

CORRECTED VERSION

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
20 September 2001 (20.09.2001)

PCT

(10) International Publication Number  
**WO 01/068857 A3**

(51) International Patent Classification<sup>7</sup>: **C12N 15/12**,  
15/62, 15/63, C07K 14/705, 16/28, C12Q 1/68, G01N  
33/577, 33/68

(21) International Application Number: PCT/US01/08329

(22) International Filing Date: 15 March 2001 (15.03.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
09/525,420 15 March 2000 (15.03.2000) US

(63) Related by continuation (CON) or continuation-in-part  
(CIP) to earlier application:  
US 09/525,420 (CIP)  
Filed on 15 March 2000 (15.03.2000)

(71) Applicant (for all designated States except US): **MIL-  
LENNIUM PHARMACEUTICALS, INC.** [US/US]; 75  
Sidney Street, Cambridge, MA 02139 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **CURTIS, Rory,  
A., J.** [US/US]; 31 Constitution Drive, Southborough,  
MA 01772 (US). **COOK, William, James** [US/US]; 22  
Reynolds Avenue, Natick, MA 01760 (US).

(74) Agents: **MANDRAGOURAS, Amy, E.**; Lahive & Cock-  
field, LLP, 28 State Street, Boston, MA 02109 et al. (US).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU,  
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,  
CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM,  
HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK,  
LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX,  
MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL,  
TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM,  
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian  
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European  
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,  
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,  
CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:  
— with international search report

(88) Date of publication of the international search report:  
30 May 2002

(48) Date of publication of this corrected version:  
10 July 2003

(15) Information about Correction:  
see PCT Gazette No. 28/2003 of 10 July 2003, Section II

*For two-letter codes and other abbreviations, refer to the "Guid-  
ance Notes on Codes and Abbreviations" appearing at the begin-  
ning of each regular issue of the PCT Gazette.*

(54) Title: 18615 AND 48003, HUMAN ION CHANNELS AND USES THEREFOR

(57) Abstract: The invention provides isolated nucleic acids molecules, designated VR-3 and VR-5 nucleic acid molecules, which are novel molecules which are members of the ion channel, e.g., calcium channel and/or vanilloid receptor, family. The invention also provides antisense nucleic acid molecules, recombinant expression vectors containing VR-3 or VR-5 nucleic acid molecules, host cells into which the expression vectors have been introduced, and nonhuman transgenic animals in which a VR-3 or VR-5 gene has been introduced or disrupted. The invention still further provides isolated VR-3 or VR-5 proteins, fusion proteins, antigenic peptides, and anti-VR-3 or anti-VR-5 antibodies. Diagnosis methods utilizing compositions of the invention are also provided.



WO 01/068857 A3

**18615 AND 48003, HUMAN ION CHANNELS AND USES THEREFOR****Related Applications**

The present application claims priority to prior filed U.S. Patent Application  
5 Serial No. 09/525,420 entitled "18615 and 48003, Novel Human Calcium  
Channel/Vanilloid Receptors and Uses Therefor", filed March 15, 2000. The content of  
the above-referenced patent application is incorporated herein by this reference in its  
entirety.

**10 Background of the Invention**

Calcium is the most abundant cation in the human body and plays a critical role  
in many physiologic processes. It is an essential component of bone as well as a well-  
known first and second messenger in signal transduction. Intracellular calcium regulates  
cell functions such as membrane excitability, release of neurotransmitters, muscle  
15 contraction, hormonal secretion, glycogen metabolism, and cell division. Extracellular  
calcium ensures the steady supply of intracellular calcium and also has other important  
functions, for example it plays a role in cell-to-cell adhesion and blood clotting.  
Intestinal absorption is the only way that calcium may enter the body. The human  
dietary intake of calcium is normally less than 1000 mg per day, of which only 30% is  
20 absorbed. This absorption occurs throughout the small intestine both through active  
transport (vitamin D-dependent) and by passive diffusion. Calcium is excreted  
primarily through the kidneys, although 95% or more is reabsorbed (resorption). Renal  
tubular resorption of filtered calcium is mainly regulated by the parathyroid hormone.

It is essential that homeostasis of calcium levels is precisely controlled. About  
25 0.1% of the total body calcium is contained in the blood and extracellular compartment  
(van Os (1987) *Biochim. Biophys. Acta*, 906:195-222). This calcium pool is maintained  
in equilibrium with the large calcium stores controlled by the bone, kidneys, and  
intestine. It is in these tissues where the bulk of calcium flux across membranes occurs  
in response to homeostatic cues.

30 Perturbations in calcium homeostasis are features of many pathological states  
(Birge and Avioli, *Clinical Disorders of Membrane Transport Processes*, Andreoli *et al.*  
Eds., Plenum Press, New York, 1987, pp. 121-140). For example, in osteoporosis,

increased resorption of bone elevates serum calcium levels, which in turn depresses the activity of parathyroid hormone. This has the effect of decreasing renal tubular resorption of calcium, which causes net urinary loss of total calcium. Other disorders associated with aberrant calcium absorption and homeostasis include intrinsic bowel  
5 disease, hepatobiliary disease, renal disease, idiopathic hypercalciuric syndromes, hypoparathyroidism, hyperthyroidism and central nervous system (CNS) disorders such as those involving neurotransmitter release (*e.g.*, Alzheimer's and Parkinson's disease).

The TRP channel family is one of the best characterized members of the capacitative calcium channel group. These channels include transient receptor potential  
10 protein and homologues thereof (to date, seven homologs and splice variants have been identified in a variety of organisms), the vanilloid receptors (also known as the capsaicin receptors), stretch-inhibitable non-selective cation channel (SIC), olfactory, mechanosensitive channel, insulin-like growth factor I-regulated calcium channel, and vitamin D-responsive apical, epithelial calcium channel (ECaC), melastatin, and the  
15 polycystic kidney disease protein family (see, *e.g.*, Montell and Rubin (1989) *Neuron* 2:1313-1323; Caterina *et al.* (1997) *Nature* 389: 816-824; Suzuki *et al.* (1999) *J. Biol. Chem.* 274: 6330-6335; Kiselyov *et al.* (1998) *Nature* 396: 478-482; Hoenderop *et al.* (1999) *J. Biol. Chem.* 274: 8375-8378; and Chen *et al.* (1999) *Nature* 401(6751): 383-6). Each of these molecules is 700 or more amino acids in length (TRP and TRP homologs  
20 have 1300 or more amino acid residues), and shares certain conserved structural features. Predominant among these structural features are six transmembrane domains, with an additional hydrophobic loop present between the fifth and sixth transmembrane domains. It is believed that this loop is integral to the activity of the pore of the channel formed upon membrane insertion (Hardie and Minke (1993) *Trends Neurosci* 16: 371-  
25 376). TRP channel proteins also include one or more ankyrin domains and frequently display a proline-rich region at the N-terminus. Although found in disparate tissues and organisms, members of the TRP channel protein family all serve to transduce signals by means of calcium entry into cells, particularly pain signals (see, *e.g.*, McClesky and Gold (1999) *Annu. Rev. Physiol.* 61: 835-856), light signals (Hardie and Minke, *supra*),  
30 or olfactory signals (Colbert *et al.* (1997) *J. Neurosci* 17(21): 8259-8269). Thus, this family of molecules may play important roles in sensory signal transduction in general.

Melastatin, a gene with homology to members of the TRP channel family, has been shown to be involved in cancer (*e.g.*, melanoma). Expression of melastatin is inversely correlated with melanoma aggressiveness such that melastatin expression was found to be downregulated in metastatic melanomas. Melastatin mRNA expression is also variably down-regulated in melanomas of intermediate thickness. These findings suggest that melastatin has a role as a suppressor of melanoma metastasis or an inhibitor of melanoma tumor progress and may be utilized as a marker for metastasis in patients with localized malignant melanoma (Duncan, *et al.* (1998) *Cancer Research* 58(7):1515-1520; Deeds, *et al.* (2000) *Hum Pathology* 31(11)1346-56; Enklaar *et al.* (2000) *Genomics* 67(2):179-87; Duncan *et al.* (2001) *J Clin Oncol* 19(2):568-576).

Vanilloid receptors (VRs) are non-selective cation channels that are structurally related to members of the TRP family of ion channels. These receptors have been proposed to mediate the entry of extracellular calcium into cells in response to the depletion of intracellular calcium stores. VRs are expressed in nociceptive neurons, as well as other cells types, and are activated by a variety of stimuli including noxious heat and protons. Capsaicin, which is a well-known agonist of VRs, induces pain behavior in humans and rodents. VR-1, a vanilloid receptor, was identified in rat sensory ganglia and is involved in pain signaling and nociception (Caterina M. J. *et al.*, (1997) *Nature* 389:816-824).

20

### **Summary of the Invention**

The present invention is based, at least in part, on the discovery of novel molecules which are members of the ion channel, *e.g.*, calcium channel and/or vanilloid receptor, family, referred to herein as “Vanilloid Receptor 3”, “Vanilloid Receptor 5”, “VR-3”, or “VR-5” nucleic acid and protein molecules. The VR-3 or VR-5 molecules of the present invention are useful as targets for developing modulating agents to regulate a variety of cellular processes, including cellular processes involved in the development and regulation of pain, as well as homeostasis of calcium levels. Furthermore, based on the discovery that the VR-3 or VR-5 molecules of the present invention are differentially expressed in tumors, *e.g.*, lung, ovarian, breast, prostate, colon, and Wilms tumors, compared to normal tissues, *e.g.*, normal lung, ovarian, breast, prostate, colon, and kidney tissue, respectively, these molecules may useful in the

30



diagnosis and treatment of cellular growth and proliferation disorders, *e.g.*, cancer, including, but not limited to, lung cancer, ovarian cancer, breast cancer, prostate cancer, colon cancer, or kidney cancer. Accordingly, in one aspect, this invention provides isolated nucleic acid molecules encoding VR-3 or VR-5 proteins or biologically active portions thereof, as well as nucleic acid fragments suitable as primers or hybridization probes for the detection of VR-3-encoding or VR-5-encoding nucleic acids.

In one embodiment, a VR-3 or VR-5 nucleic acid molecule of the invention is at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 83%, 85%, 86%, 90%, 91%, 92%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.5% or more identical to the nucleotide sequence (*e.g.*, to the entire length of the nucleotide sequence) shown in SEQ ID NO:1, 3, 4, or 6 or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number PTA-2013, or a complement thereof.

In a preferred embodiment, the isolated nucleic acid molecule includes the nucleotide sequence shown SEQ ID NO:1, 3, 4, or 6, or a complement thereof. In another embodiment, the nucleic acid molecule includes SEQ ID NO:3 and nucleotides 1-277 of SEQ ID NO:1. In another embodiment, the nucleic acid molecule includes SEQ ID NO:3 and nucleotides 2456-3026 of SEQ ID NO:1. In another embodiment, the nucleic acid molecule includes SEQ ID NO:6 and nucleotides 1-83 of SEQ ID NO:4. In another embodiment, the nucleic acid molecule includes SEQ ID NO:6 and nucleotides 2700-3245 of SEQ ID NO:4. In another preferred embodiment, the nucleic acid molecule consists of the nucleotide sequence shown in SEQ ID NO:1, 3, 4, or 6. In another preferred embodiment, the nucleic acid molecule includes a fragment of at least 50, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, 2100, 2200, 2242, 2300, 2400, 2500, 2600, 2700, 2800, 2900, 3000, 3100, 3200, or more nucleotides (*e.g.*, contiguous nucleotides) of the nucleotide sequence of SEQ ID NO:1, 3, 4, or 6, or a complement thereof.

In another embodiment, a VR-3 or VR-5 nucleic acid molecule includes a nucleotide sequence encoding a protein having an amino acid sequence sufficiently identical to the amino acid sequence of SEQ ID NO:2 or 5 or an amino acid sequence encoded by the DNA insert of the plasmid deposited with ATCC as Accession Number PTA-2013. In a preferred embodiment, a VR-3 or VR-5 nucleic acid molecule includes a nucleotide sequence encoding a protein having an amino acid sequence at least 50%,

55%, 60%, 65%, 70%, 75%, 80%, 85%, 86%, 90%, 91%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.5%, or more identical to the entire length of the amino acid sequence of SEQ ID NO:2 or 5, or the amino acid sequence encoded by the DNA insert of the plasmid deposited with ATCC as Accession Number PTA-2013.

5 In another preferred embodiment, an isolated nucleic acid molecule encodes the amino acid sequence of human VR-3 or VR-5. In yet another preferred embodiment, the nucleic acid molecule includes a nucleotide sequence encoding a protein having the amino acid sequence of SEQ ID NO:2 or 5, or the amino acid sequence encoded by the DNA insert of the plasmid deposited with ATCC as Accession Number PTA-2013.

10 In another preferred embodiment, the nucleic acid molecule is at least 362 nucleotides in length. In a further preferred embodiment, the nucleic acid molecule is at least 362 nucleotides in length and encodes a protein having a VR-3 activity (as described herein). In yet another preferred embodiment, the nucleic acid molecule is at least 519 nucleotides in length. In a further preferred embodiment, the nucleic acid  
15 molecule is at least 519 nucleotides in length and encodes a protein having a VR-5 activity (as described herein).

Another embodiment of the invention features nucleic acid molecules, preferably VR-3 or VR-5 nucleic acid molecules, which specifically detect VR-3 or VR-5 nucleic acid molecules relative to nucleic acid molecules encoding non-VR-3 or non-VR-5  
20 proteins. For example, in one embodiment, such a nucleic acid molecule is at least 100-500, 500-1000, 1000-1500, 1500-2000, 2000-2500, 2500-3000, 3000-3200, or more nucleotides in length and hybridizes under stringent conditions to a nucleic acid molecule comprising the nucleotide sequence shown in SEQ ID NO:1 or 4, the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as  
25 Accession Number PTA-2013, or a complement thereof.

In preferred embodiments, the nucleic acid molecules are at least 15 (*e.g.*, contiguous) nucleotides in length and hybridize under stringent conditions to nucleotides 1-65 or 2986-3026 of SEQ ID NO:1. In other preferred embodiments, the nucleic acid molecules comprise nucleotides 1-65 or 2986-3026 of SEQ ID NO:1. In other preferred  
30 embodiments, the nucleic acid molecules consist of nucleotides 1-65 or 2986-3026 of SEQ ID NO:1.

In preferred embodiments, the nucleic acid molecules are at least 15 (*e.g.*, contiguous) nucleotides in length and hybridize under stringent conditions to nucleotides 1-31 of SEQ ID NO:4. In other preferred embodiments, the nucleic acid molecules comprise nucleotides 1-31 of SEQ ID NO:4. In other preferred embodiments, the  
5 nucleic acid molecules consist of nucleotides 1-31 of SEQ ID NO:4.

In other preferred embodiments, the nucleic acid molecule encodes a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NO:2 or 5, or an amino acid sequence encoded by the DNA insert of the plasmid deposited with ATCC as Accession Number PTA-2013, wherein the nucleic acid  
10 molecule hybridizes to a nucleic acid molecule comprising SEQ ID NO:1, 3, 4 or 6 under stringent conditions.

Another embodiment of the invention provides an isolated nucleic acid molecule which is antisense to a VR-3 or VR-5 nucleic acid molecule, *e.g.*, the coding strand of a VR-3 or VR-5 nucleic acid molecule.

15 Another aspect of the invention provides a vector comprising a VR-3 or VR-5 nucleic acid molecule. In certain embodiments, the vector is a recombinant expression vector. In another embodiment, the invention provides a host cell containing a vector of the invention. In yet another embodiment, the invention provides a host cell containing a nucleic acid molecule of the invention. The invention also provides a method for  
20 producing a protein, preferably a VR-3 or VR-5 protein family member, by culturing a host cell in a suitable medium, *e.g.*, a mammalian host cell such as a non-human mammalian cell, of the invention containing a recombinant expression vector, such that the protein is produced.

Another aspect of this invention features isolated or recombinant VR-3 proteins  
25 and polypeptides. In preferred embodiments, the isolated VR-3 protein family member includes at least one or more of the following domains: an ankyrin repeat domain, a pore domain, and/or a transmembrane domain.

Another aspect of this invention features isolated or recombinant VR-5 proteins and polypeptides. In preferred embodiments, the isolated VR-5 protein family member  
30 includes at least one or more of the following domains: an ankyrin repeat domain, a pore domain, a transmembrane domains, and/or an ion transport protein domain.

In a preferred embodiment, the VR-3 protein family member has an amino acid sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 86%, 90%, 91%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.5%, or more identical to the amino acid sequence of SEQ ID NO:2, and includes at least one or more of the following domains:

5 an ankyrin repeat domain, a pore domain, and/or a transmembrane domain.

In another preferred embodiment, the VR-5 protein family member has an amino acid sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 86%, 90%, 91%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.5%, or more identical to the amino acid sequence of SEQ ID NO:5, or the amino acid sequence encoded by the DNA insert of the plasmid deposited with ATCC as Accession Number PTA-2013, and includes at least one or more of the following domains: an ankyrin repeat domain, a pore domain, a transmembrane domains, and/or an ion transport protein domain.

In another preferred embodiment, the VR-3 protein family member plays a role in calcium homeostasis, pain signaling, and/or cellular growth and/or proliferation and includes at least one or more of the following domains: an ankyrin repeat domain, a pore domain, and/or a transmembrane domain.

In another preferred embodiment, the VR-5 protein family member plays a role in calcium homeostasis, pain signaling, and/or cellular growth and/or proliferation and includes at least one or more of the following domains: an ankyrin repeat domain, a pore domain, a transmembrane domains, and/or an ion transport protein domains.

In yet another preferred embodiment, the VR-3 protein family member is encoded by a nucleic acid molecule having a nucleotide sequence which hybridizes under stringent hybridization conditions to a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1 or 3, and includes at least one or more of the following domains: an ankyrin repeat domain, a pore domain, and/or a transmembrane domain.

In yet another preferred embodiment, the VR-5 protein family member is encoded by a nucleic acid molecule having a nucleotide sequence which hybridizes under stringent hybridization conditions to a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:4 or 6, and includes at least one or more of the following domains: an ankyrin repeat domain, a pore domain, a transmembrane domains, and/or an ion transport protein domain.

In another embodiment, the invention features fragments of the protein having the amino acid sequence of SEQ ID NO:2 or 5, wherein the fragment comprises at least 15, 20, 30, 40, 50, 60, 70, 80, 90, or 100 amino acids (*e.g.*, contiguous amino acids) of the amino acid sequence of SEQ ID NO:2 or 5, or an amino acid sequence encoded by the DNA insert of the plasmid deposited with the ATCC as Accession Number PTA-2013. In another embodiment, the protein, preferably a VR-3 or VR-5 protein, has the amino acid sequence of SEQ ID NO:2 or 5.

In another embodiment, the invention features an isolated VR-3 or VR-5 protein family member which is encoded by a nucleic acid molecule consisting of a nucleotide sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 82.6%, 85%, 86%, 90%, 95%, 96%, 97%, 98%, 99%, 99.1% or more identical to a nucleotide sequence of SEQ ID NO:1, 3, 4 or 6, or a complement thereof. This invention further features an isolated protein, preferably a VR-3 or VR-5 protein, which is encoded by a nucleic acid molecule consisting of a nucleotide sequence which hybridizes under stringent hybridization conditions to a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1, 3, 4 or 6, or a complement thereof.

The proteins of the present invention or portions thereof, *e.g.*, biologically active portions thereof, can be operatively linked to a non-VR-3 or a non-VR-5 polypeptide (*e.g.*, heterologous amino acid sequences) to form fusion proteins. The invention further features antibodies, such as monoclonal or polyclonal antibodies, that specifically bind proteins of the invention, preferably VR-3 or VR-5 proteins. In addition, the VR-3 or VR-5 proteins or biologically active portions thereof can be incorporated into pharmaceutical compositions, which optionally include pharmaceutically acceptable carriers.

In another aspect, the present invention provides a method for detecting the presence of a VR-3 or VR-5 nucleic acid molecule, protein or polypeptide in a biological sample by contacting the biological sample with an agent capable of detecting a VR-3 or VR-5 nucleic acid molecule, protein or polypeptide such that the presence of a VR-3 or VR-5 nucleic acid molecule, protein or polypeptide is detected in the biological sample.

In another aspect, the present invention provides a method for detecting the presence of VR-3 or VR-5 activity in a biological sample by contacting the biological sample with an agent capable of detecting an indicator of VR-3 or VR-5 activity such that the presence of VR-3 or VR-5 activity is detected in the biological sample.

5 In another aspect, the invention provides a method for modulating VR-3 or VR-5 activity comprising contacting a cell capable of expressing VR-3 or VR-5 with an agent that modulates VR-3 or VR-5 activity such that VR-3 or VR-5 activity in the cell is modulated. In one embodiment, the agent inhibits VR-3 or VR-5 activity. In another embodiment, the agent stimulates VR-3 or VR-5 activity. In one embodiment, the agent  
10 is an antibody that specifically binds to a VR-3 or VR-5 protein. In another embodiment, the agent modulates expression of VR-3 or VR-5 by modulating transcription of a VR-3 or VR-5 gene or translation of a VR-3 or VR-5 mRNA. In yet another embodiment, the agent is a nucleic acid molecule having a nucleotide sequence that is antisense to the coding strand of a VR-3 or VR-5 mRNA or a VR-3 or VR-5  
15 gene.

In one embodiment, the methods of the present invention are used to treat a subject having a disorder characterized by aberrant or unwanted VR-3 or VR-5 protein or nucleic acid expression or activity by administering an agent which is a VR-3 or VR-5 modulator to the subject. In one embodiment, the VR-3 or VR-5 modulator is a VR-3  
20 or VR-5 protein. In another embodiment the VR-3 or VR-5 modulator is a VR-3 or VR-5 nucleic acid molecule. In yet another embodiment, the VR-3 or VR-5 modulator is an antibody, ribozyme, peptide, peptidomimetic, antisense oligonucleotide, or other small molecule. In a preferred embodiment, the disorder characterized by aberrant or unwanted VR-3 or VR-5 protein or nucleic acid expression is a calcium homeostasis  
25 related disorder. In another preferred embodiment, the disorder characterized by aberrant or unwanted VR-3 or VR-5 protein or nucleic acid expression is cancer, *e.g.*, lung cancer, ovarian cancer, breast cancer, prostate cancer, colon cancer, or Wilms tumors. In yet another preferred embodiment, the disorder characterized by aberrant or unwanted VR-3 or VR-5 protein or nucleic acid expression is pain or a pain disorder.

30 The present invention also provides diagnostic assays for identifying the presence or absence of a genetic alteration characterized by at least one of (i) aberrant modification or mutation of a gene encoding a VR-3 or VR-5 protein; (ii) mis-regulation

of the VR-3 or VR-5 gene; and (iii) aberrant post-translational modification of a VR-3 or VR-5 protein, wherein a wild-type form of the gene encodes a protein with a VR-3 or VR-5 activity.

In another aspect the invention provides methods for identifying a compound  
5 that binds to or modulates the activity of a VR-3 or VR-5 protein, by providing an indicator composition comprising a VR-3 or VR-5 protein having VR-3 or VR-5 activity, contacting the indicator composition with a test compound, and determining the effect of the test compound on VR-3 or VR-5 activity in the indicator composition to identify a compound that modulates the activity of a VR-3 or VR-5 protein.

10

Other features and advantages of the invention will be apparent from the following detailed description and claims.

#### **Brief Description of the Drawings**

15 *Figure 1* depicts the cDNA sequence and predicted amino acid sequence of the human VR-3. The nucleotide sequence corresponds to nucleic acids 1 to 3026 of SEQ ID NO:1. The amino acid sequence corresponds to amino acids 1 to 725 of SEQ ID NO:2. The coding region of the human VR-3 is shown in SEQ ID NO:3.

20 *Figure 2* depicts the cDNA sequence and predicted amino acid sequence of the human VR-5. The nucleotide sequence corresponds to nucleic acids 1 to 3245 of SEQ ID NO:4. The amino acid sequence corresponds to amino acids 1 to 871 of SEQ ID NO:5. The coding region of the human VR-5 is shown in SEQ ID NO:6.

25 *Figure 3* depicts an alignment of the amino acid sequence of human VR-3 with the rat calcium transporter (GenBank Accession No. AF160798) and the rabbit epithelial calcium channel (GenBank Accession No. AJ133128) using the CLUSTALW (1.74) multiple sequence alignment program.

*Figure 4* depicts an alignment of the amino acid sequence of human VR-5 with the amino acid sequence of the *Mus musculus* ion channel (GenBank Accession No. AB021875) using the GAP program in the GCG software package (Blosum 62 matrix)  
30 and a gap weight of 12 and a length weight of 4.

*Figure 5* depicts an alignment of the nucleotide sequence of human VR-5 with the nucleotide sequence of the *Mus musculus* ion channel (GenBank Accession Number AB021875) using the CLUSTALW (1.74) multiple sequence alignment program.

5        *Figure 6* depicts a structural, hydrophobicity, and antigenicity analysis of the human VR-3 protein.

*Figure 7* depicts a structural, hydrophobicity, and antigenicity analysis of the human VR-5 protein.

*Figure 8* depicts the results of a search which was performed against the HMM  
10 database using the amino acid sequence of human VR-3. This search resulted in the identification of three “ankyrin repeat” domains domain in the human VR-3 protein.

*Figure 9* depicts the results of a search which was performed against the HMM database using the amino acid sequence of human VR-5. This search resulted in the identification of three “ankyrin repeat” domains and one “ion transport protein ” domain  
15 in the human VR-5 protein.

*Figure 10* depicts the results of a search performed against the HMM database using the amino acid sequence of human VR-3. This search resulted in the local alignment of the human VR-3 protein with the rat VR-1 protein, the protein olfactory channel/VR, and the transmembrane calcium receptor/ion transport protein.

20        *Figure 11* depicts the results of a search performed against the HMM database using the amino acid sequence of human VR-5. This search resulted in the local alignment of the human VR-5 protein with the rat VR-1 protein, the protein olfactory channel/VR, and the transmembrane calcium receptor/ion transport protein.

## 25        **Detailed Description of the Invention**

The present invention is based, at least in part, on the discovery of novel molecules which are members of the ion channel, *e.g.*, calcium channel and/or vanilloid receptor, family. Described herein is the isolation of two human ion channels, *e.g.*, calcium channel/vanilloid receptors, referred to herein as “Vanilloid Receptor-3” or  
30 “VR-3” or “VR-5” and as “Vanilloid Receptor 5” or “VR-5.”



The VR-3 and VR-5 sequences of the present invention are similar to that of rat VR-1. VR-1 is a vanilloid gated, non-selective cation channel which resembles members of the transient receptor potential (TRP) ion channel family (described in Montell *et al.* (1989) *Neuron* 2:1313-1323) that mediate the influx of extracellular calcium in response to depletion of intracellular calcium stores. Hydrophilicity analysis has indicated that rat VR-1 contains six transmembrane domains (predicted to be mostly  $\alpha$ -helices). The amino terminal hydrophilic segment contains three ankyrin repeat domains. The rat VR-1 was identified in rat sensory ganglia (Caterina M. J. *et al.*, (1997) *Nature* 389:816-824). It has been shown that VR-1 knockout mice are impaired in their detection of painful heat, exhibit no vanilloid-evoked pain behavior, and show little thermal hypersensitivity after inflammation (Szallasi and Blumberg (1999) *Pharmacol. Rev.* 51:159-211; Tominaga, *et al.* (1998) *Neuron* 21:531; Caterina *et al.* (2000) *Science* 288:306). Based on homology to VR-1 and the discovery that VR-3 and VR-5 are expressed in brain (*e.g.*, cortex and hypothalamus), and spinal cord, VR-3 and VR-5 may be involved in nociception (*e.g.*, chemical, mechanical, or thermal nociception) and thereby may modulate pain elicitation. Accordingly, the VR-3 and VR-5 molecules of the present invention act as targets for developing novel diagnostic targets and therapeutic agents to control pain and pain disorders.

As used herein, an "ion channel" includes a protein or polypeptide which is involved in receiving, conducting, and transmitting signals in an electrically excitable cell, *e.g.*, a neuronal or muscle cell. Ion channels include vanilloid receptors, calcium channels, potassium channels, and sodium channels. The VR-3 and VR-5 molecules of the present invention are highly expressed in kidney, indicating that these molecules may function as calcium channels. As used herein, a "calcium channel" includes a protein or polypeptide which is involved in receiving, conducting, and transmitting calcium ion-based signals in an electrically excitable cell. Calcium channels are calcium ion selective, and can determine membrane excitability (the ability of, for example, a neuronal cell to respond to a stimulus and to convert it into a sensory impulse). Calcium channels can also influence the resting potential of membranes, wave forms and frequencies of action potentials, and thresholds of excitation. Calcium channels are typically expressed in electrically excitable cells, *e.g.*, neuronal cells, and may form heteromultimeric structures (*e.g.*, composed of more than one type of subunit). Calcium

channels may also be found in non-excitabile cells (*e.g.*, adipose cells or liver cells), where they may play a role in, *e.g.*, signal transduction. Examples of calcium channels include the low-voltage-gated channels and the high-voltage-gated channels. Calcium channels are described in, for example, Davila *et al.* (1999) *Annals New York Academy of Sciences* 868:102-17 and McEnery, M.W. *et al.* (1998) *J. Bioenergetics and Biomembranes* 30(4): 409-418, the contents of which are incorporated herein by reference.

As used herein, a “vanilloid receptor” includes a non-selective cation channel that is structurally related to the TRP family of ion channels. Vanilloid receptors are also known as capsaicin receptors. Vanilloid receptors share several physical characteristics including an N-terminal cytoplasmic domain which contains three ankyrin repeats, six transmembrane domains, a pore-loop region located between transmembrane domains 5 and 6, and several kinase consensus sequences. Members of the vanilloid receptor (VR) family have been proposed to mediate the entry of extracellular calcium into cells, *e.g.*, in response to the depletion of intracellular calcium stores. VRs are typically expressed in nociceptive neurons among other cells types and are directly activated by harmful heat, extracellular protons, and vanilloid compounds. VRs may also be expressed in nonsensory tissues and may mediate inflammatory rather than acute thermal pain. Vanilloid receptors are described in, for example, Caterina, M.J. (1997) *Nature* 389:816-824 and Caterina, M.J. (2000) *Science* 288:306-313) the contents of which are incorporated herein by reference. As the VR-3 and VR-5 molecules of the present invention may modulate ion channel mediated activities (*e.g.*, calcium channel- and/or vanilloid receptor- mediated activities), they may be useful for developing novel diagnostic and therapeutic agents for ion channel associated disorders (*e.g.*, calcium channel and/or vanilloid receptor associated disorders).

As used herein, an “ion channel associated disorder” includes a disorder, disease or condition which is characterized by a misregulation of ion channel (*e.g.*, calcium channel) and/or vanilloid receptor mediated activity. For example, a “calcium channel associated disorder” includes a disorder, disease or condition which is characterized by a misregulation of calcium channel mediated activity. A “vanilloid receptor associated disorder” includes a disorder, disease or condition which is characterized by a misregulation of vanilloid receptor mediated activity. Ion channel associated disorders,

*e.g.*, calcium channel associated disorders and/or vanilloid receptor associated disorders, include CNS disorders, such as cognitive and neurodegenerative disorders, examples of which include, but are not limited to, Alzheimer's disease, dementias related to Alzheimer's disease (such as Pick's disease), Parkinson's and other Lewy diffuse body diseases, senile dementia, Huntington's disease, Gilles de la Tourette's syndrome, multiple sclerosis, amyotrophic lateral sclerosis, progressive supranuclear palsy, epilepsy, Jakob-Creutzfeldt disease, or AIDS related dementia; autonomic function disorders such as hypertension and sleep disorders, and neuropsychiatric disorders, such as depression, schizophrenia, schizoaffective disorder, korsakoff's psychosis, mania, anxiety disorders, or phobic disorders; learning or memory disorders, *e.g.*, amnesia or age-related memory loss, attention deficit disorder, psychoactive substance use disorders, anxiety, phobias, panic disorder, as well as bipolar affective disorder, *e.g.*, severe bipolar affective (mood) disorder (BP-1), and bipolar affective neurological disorders, *e.g.*, migraine and obesity. Further CNS-related disorders include, for example, those listed in the American Psychiatric Association's Diagnostic and Statistical manual of Mental Disorders (DSM), the most current version of which is incorporated herein by reference in its entirety.

Ion channel associated disorders, *e.g.*, calcium channel disorders and/or vanilloid receptor associated disorders, also include pain disorders. As used herein, the term "pain disorder" includes a disorder affecting pain signaling mechanisms. Pain disorders include disorders characterized by aberrant (*e.g.*, excessive or amplified) pain. The VR-3 or VR-5 molecules may provide novel diagnostic targets and therapeutic agents to control pain in a variety of disorders, diseases, or conditions which are characterized by a deregulated, *e.g.*, upregulated or downregulated, pain response. For example, VR-3 or VR-5 molecules may provide novel diagnostic targets and therapeutic agents to control the exaggerated pain response elicited during various forms of tissue injury, *e.g.*, inflammation, infection, and ischemia, usually referred to as hyperalgesia (described in, for example, Fields, H.L. (1987) *Pain*, New York; McGraw-Hill). Further examples of pain and/or pain disorders include posttherapeutic neuralgia, diabetic neuropathy, postmastectomy pain syndrome, stump pain, reflex sympathetic dystrophy, trigeminal neuralgia, neuropathic pain, orofacial neuropathic pain, osteoarthritis, rheumatoid arthritis, fibromyalgia syndrome, tension myalgia, Guillian-Barre syndrome, Meralgia

paraesthetica, burning mouth syndrome, fibrocitis, myofascial pain syndrome, idiopathic pain disorder, temporomandibular joint syndrome, atypical odontalgia, loin pain, haematuria syndrome, non-cardiac chest pain, low back pain, chronic nonspecific pain, pain associated with surgery, psychogenic pain, tooth pain, musculoskeletal pain disorder, chronic pelvic pain, nonorganic chronic headache, tension-type headache, 5 cluster headache, migraine, complex regional pain syndrome, vaginismus, nerve trunk pain, somatoform pain disorder, cyclical mastalgia, chronic fatigue syndrome, multiple somatization syndrome, chronic pain disorder, somatization disorder, Syndrome X, facial pain, idiopathic pain disorder, posttraumatic rheumatic pain modulation disorder 10 (fibrositis syndrome), hyperalgesia, and Tangier disease.

As used herein, the term "pain signaling mechanisms" includes the cellular mechanisms involved in the development and regulation of pain, *e.g.*, pain elicited by noxious chemical, mechanical, or thermal stimuli, in a subject, *e.g.*, a mammal such as a human. In mammals, the initial detection of noxious chemical, mechanical, or thermal 15 stimuli, a process referred to as "nociception", occurs predominantly at the peripheral terminals of specialized, small diameter primary afferent neurons called polymodal nociceptors. These afferent neurons transmit the information to the central nervous system, evoking a perception of pain or discomfort and initiating appropriate protective reflexes. Vanilloid receptors, *e.g.*, the VR-3 and VR-5 molecules of the present 20 invention, present on these afferent neurons, may be involved in detecting these noxious chemical, mechanical, or thermal stimuli and transducing this information into membrane depolarization events. Thus, the VR-3 and VR-5 molecules, by participating in pain signaling mechanisms, may modulate pain elicitation and provide novel diagnostic targets and therapeutic agents to control pain and pain disorders.

25 The VR-3 or VR-5 molecules of the present invention also play a role in calcium homeostasis. As used herein, the term "calcium homeostasis" includes cellular mechanisms involved in maintaining an equilibrium of intracellular or extracellular calcium concentration. Such mechanisms include the movement of calcium ions across cellular membranes (*e.g.*, intestine or kidney cellular membranes) in response to 30 biological cues. The maintenance of calcium homeostasis is particularly important for an organism's nutritional needs. Important calcium transport processes are known to occur in the intestine and in the kidney. Thus, the VR-3 and VR-5 molecules, by

participating in calcium homeostasis mechanisms, can modulate calcium homeostasis mechanisms and provide novel diagnostic targets and therapeutic agents to control calcium homeostasis related disorders.

As used herein, the term “calcium homeostasis related disorders” includes disorders which are characterized by aberrant, *e.g.*, upregulated or downregulated, extracellular or intracellular calcium concentrations. Examples of such disorders include idiopathic hypercalciuria, sarcoidosis and other granulomatous disorders, primary hyperparathyroidism, diabetes, phosphorus depletion, osteoporosis, intrinsic bowel disease, hepatobiliary disease, renal disease, hyperthyroidism, and hypoparathyroidism, and CNS disorders, *e.g.*, Alzheimer’s disease or Parkinson’s disease.

The present invention is also based, at least in part, on the discovery that the VR-3 and VR-5 molecules are differentially expressed in tumors. VR-3 is differentially expressed in breast, colon, and prostate tumors as compared to normal breast, colon and prostate tissues. VR-5 is differentially expressed in lung, ovary, breast, and Wilms tumors, as compared to normal lung, ovary, breast, and kidney tissue. Accordingly, the VR-3 and VR-5 molecules of the present invention provide novel diagnostic targets and therapeutic agents to control cellular growth and/or proliferation disorders, *e.g.*, cancer.

As used herein, a “cellular growth and/or proliferation disorder” includes a disease or disorder that affects a cell growth or proliferation process. As used herein, a “cellular growth or proliferation process” is a process by which a cell increases in number, size or content, by which a cell develops a specialized set of characteristics which differ from that of other cells, or by which a cell moves closer to or further from a particular location or stimulus. A cellular growth or proliferation process includes the metabolic processes of the cell and cellular transcriptional activation mechanisms. A cellular growth or proliferation disorder may be characterized by aberrantly regulated cell growth, proliferation, differentiation, or migration. Cellular growth or proliferation disorders include tumorigenic disease or disorders. As used herein, a “tumorigenic disease or disorder” includes a disease or disorder characterized by aberrantly regulated cell growth, proliferation, differentiation, adhesion, or migration, resulting in the production of or tendency to produce tumors. As used herein, a “tumor” includes a normal benign or malignant mass of tissue. Examples of cellular growth or proliferation

disorders include, but are not limited to, cancer, *e.g.*, carcinoma, sarcoma, or leukemia, examples of which include, but are not limited to, colon, ovarian, lung, breast, endometrial, uterine, hepatic, gastrointestinal, prostate, and brain cancer; Wilms tumors; tumorigenesis and metastasis; skeletal dysplasia; and hematopoietic and/or  
5 myeloproliferative disorders.

“Differential expression”, as used herein, includes both quantitative as well as qualitative differences in the temporal and/or tissue expression pattern of a gene. Thus, a differentially expressed gene may have its expression activated or inactivated in normal versus cellular growth or proliferation disease states. The degree to which  
10 expression differs in normal versus cellular growth or proliferation disease states or control versus experimental states need only be large enough to be visualized via standard characterization techniques, *e.g.*, quantitative PCR, Northern analysis, or subtractive hybridization. The expression pattern of a differentially expressed gene may be used as part of a prognostic or diagnostic cellular growth or proliferation disorder  
15 evaluation, or may be used in methods for identifying compounds useful for the treatment of cellular growth or proliferation disorder. In addition, a differentially expressed gene involved in tumorigenic disorders may represent a target gene such that modulation of the expression level of this gene or the activity of the gene product may act to ameliorate a cellular growth or proliferation disorder. Compounds that modulate  
20 target gene expression or activity of the target gene product can be used in the treatment of cellular growth or proliferation disorders. Although the VR-3 and VR-5 genes described herein may be differentially expressed with respect to cellular growth or proliferation disorders, and/or their products may interact with gene products important to cellular growth or proliferation disorders, the genes may also be involved in  
25 mechanisms important to additional tumor cell processes.

The term "family" when referring to the protein and nucleic acid molecules of the invention is intended to mean two or more proteins or nucleic acid molecules having a common structural domain or motif and having sufficient amino acid or nucleotide sequence homology as defined herein. Such family members can be naturally or non-  
30 naturally occurring and can be from either the same or different species. For example, a family can contain a first protein of human origin, as well as other, distinct proteins of

human origin or alternatively, can contain homologues of non-human origin. Members of a family may also have common functional characteristics.

For example, the family of VR-3 or VR-5 proteins comprise at least one, and preferably five to six "transmembrane domains." As used herein, the term

5 "transmembrane domain" includes an amino acid sequence of about 15 amino acid residues in length which spans the plasma membrane. More preferably, a transmembrane domain includes about at least 10, 15, 20, 25, 30, 35, 40, 45 or more amino acid residues and spans the plasma membrane. Transmembrane domains are rich in hydrophobic residues, and typically have a helical structure. In a embodiment, at  
10 least 50%, 60%, 70%, 80%, 90%, 95% or more of the amino acid residues of a transmembrane domain are hydrophobic, *e.g.*, leucines, isoleucines, tyrosines, or tryptophans. Transmembrane domains are described in, for example, Zagotta W.N. et al, (1996) *Annual Rev. Neurosci.* 19: 235-63, the contents of which are incorporated herein by reference. Amino acid residues 328-349, 386-402, 420-442, 456-482, 493-  
15 512, and 553-577 of the human VR-3 polypeptide (SEQ ID NO:2) comprise transmembrane domains (Figure 6). Amino acid residues 466-490, 511-529, 551-568, 575-606, 617-636, and 693-717 of the human VR-5 polypeptide (SEQ ID NO:5) also comprise transmembrane domains (Figure 7).

In another embodiment, a VR-3 or VR-5 molecule of the present invention is  
20 identified based on the presence of an "ankyrin repeat domain" in the protein or corresponding nucleic acid molecule. As used herein, the term "ankyrin repeat domain" includes a protein domain having an amino acid sequence of about 30-50 amino acid residues and having a bit score for the alignment of the sequence to the ankyrin repeat domain (HMM) of at least 6. Preferably, an ankyrin repeat domain includes at least  
25 about 30-45, more preferably about 30-40 amino acid residues, or about 30-38 amino acids and has a bit score for the alignment of the sequence to the ankyrin repeat domain (HMM) of at least 2, 5-10, 10-20, 20-30, 30-40, 40-60 or greater. The ankyrin repeat domain HMM has been assigned the PFAM Accession PF00023 (<http://genome.wustl.edu/Pfam/.html>). Ankyrin repeats are involved in protein-protein  
30 interactions and are described in, for example, Ketchum K.A., *et al.* (1996) *FEBS Letters* 378:19-26, the contents of which are incorporated herein by reference.

In another embodiment, a VR-3 or VR-5 molecule of the present invention is identified based on the presence of at least one "pore domain" between the fifth and sixth transmembrane domains. As used herein, the term "pore domain" includes an overall hydrophobic amino acid sequence which is located between two transmembrane domains of a calcium channel protein, preferably transmembrane domains 5 and 6, and which is believed to be a major determinant of ion selectivity and channel activity in calcium channels. Pore domains are described in, for example, Vannier *et al.* (1998) *J. Biol. Chem.* 273: 8675-8679 and Phillips, A. M. *et al.* (1992) *Neuron* 8, 631-642, the contents of which are incorporated herein by reference. VR-3 or VR-5 molecules having at least one pore domain are within the scope of the invention. Amino acid residues 523-544 of the human VR-3 polypeptide (SEQ ID NO:2) comprise a pore domain. Amino acid residues 666-683 of the human VR-5 polypeptide (SEQ ID NO:5) also comprise a pore domain.

In another embodiment, a VR-5 molecule of the present invention is identified based on the presence of an "ion transport protein domain." As used herein, the term "ion transport protein domain" includes a protein domain having an amino acid sequence of at least about 200-300, more preferably at least about 220-280 or at least about 235-260 amino acid residues and having a bit score for the alignment of the sequence to the ion transport protein domain (HMM) of at least about 1, 5, 10, 20, 30, 40, 50 or greater. The ion transport protein domain HMM has been assigned the PFAM Accession Number PF00520 (<http://genome.wustl.edu/Pfam/.html>). Proteins exhibiting this domain include sodium, potassium, and calcium ion channels.

To identify the presence of an ankyrin repeat domain or an ion transport protein domain in a VR-3 or VR-5 protein and make the determination that a protein of interest has a particular profile, the amino acid sequence of the protein is searched against a database of HMMs (*e.g.*, the Pfam database, release 2.1) using the default parameters ([http://www.sanger.ac.uk/Software/Pfam/HMM\\_search](http://www.sanger.ac.uk/Software/Pfam/HMM_search)). A description of the Pfam database can be found in Sonhammer *et al.* (1997) *Proteins* 28(3)405-420 and a detailed description of HMMs can be found, for example, in Gribskov *et al.* (1990) *Meth. Enzymol.* 183:146-159; Gribskov *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:4355-4358; Krogh *et al.* (1994) *J. Mol. Biol.* 235:1501-1531; and Stultz *et al.* (1993) *Protein Sci.* 2:305-314, the contents of which are incorporated herein by reference. A search



was performed against the HMM database resulting in the identification of three ankyrin repeat domains in the amino acid sequence of SEQ ID NO:2 (at about residues 78-108, 116-148, and 162-194). The search also identified the presence of three ankyrin repeat domains in SEQ ID NO:5 (at about residues 237-269, 284-319, and 369-400). The search further identified an ion transport protein domain in the amino acid sequence of SEQ ID NO:5 (at about residues 473-718). The results of this search are set forth in Figures 8 and 9.

Isolated VR-3 or VR-5 proteins of the present invention, have an amino acid sequence sufficiently identical to the amino acid sequence of SEQ ID NO:2 or 5, or are encoded by a nucleotide sequence sufficiently identical to SEQ ID NO:1, 3, 4 or 6. As used herein, the term "sufficiently identical" refers to a first amino acid or nucleotide sequence which contains a sufficient or minimum number of identical or equivalent (*e.g.*, an amino acid residue which has a similar side chain) amino acid residues or nucleotides to a second amino acid or nucleotide sequence such that the first and second amino acid or nucleotide sequences share common structural domains or motifs and/or a common functional activity. For example, amino acid or nucleotide sequences which share common structural domains have at least 30%, 40%, or 50% homology, preferably 60% homology, more preferably 70%-80%, and even more preferably 90-95% homology across the amino acid sequences of the domains and contain at least one and preferably two structural domains or motifs, are defined herein as sufficiently identical. Furthermore, amino acid or nucleotide sequences which share at least 30%, 40%, or 50%, preferably 60%, more preferably 70-80%, or 90-95% homology and share a common functional activity are defined herein as sufficiently identical.

As used interchangeably herein, a "VR-3 or VR-5 activity", "VR-3 activity", "VR-5 activity", "biological activity of VR-3", "biological activity of VR-5", "functional activity of VR-3", or "functional activity of VR-5", includes an activity exerted by a VR-3 or VR-5 protein, polypeptide or nucleic acid molecule on a VR-3- or VR-5-responsive cell or tissue, or on a VR-3 or VR-5 protein substrate, as determined *in vivo*, or *in vitro*, according to standard techniques. In one embodiment, a VR-3 or VR-5 activity is a direct activity, such as an association with a VR-3- or VR-5-target molecule. As used herein, a "target molecule" or "binding partner" is a molecule with which a VR-3 or VR-5 protein binds or interacts in nature, such that VR-3- or VR-5-mediated

function is achieved. A VR-3 or VR-5 target molecule can be a non-VR-3 or non-VR-5 molecule or a VR-3 or VR-5 protein or polypeptide of the present invention. In an exemplary embodiment, a VR-3 or VR-5 target molecule is a VR-3 or VR-5 ligand, *e.g.*, a vanilloid molecule or a vanilloid-containing compound such as capsaicin.

5 Alternatively, a VR-3 or VR-5 activity is an indirect activity, such as a cellular signaling activity mediated by interaction of the VR-3 or VR-5 protein with a VR-3 or VR-5 ligand, *e.g.*, a vanilloid or a vanilloid-containing compound such as capsaicin.

Preferably, a VR-3 or VR-5 activity is the ability to modulate the transmission of pain via, *e.g.*, pain signaling mechanisms. Also preferably, a VR-3 or VR-5 activity is the  
10 ability to modulate the transport of calcium via, *e.g.*, calcium signaling mechanisms. In addition, a VR-3 or VR-5 activity also includes the modulation of cellular growth and/or proliferation and/or tumorigenesis.

Accordingly, another embodiment of the invention features isolated VR-3 or VR-5 polypeptides having a VR-3 or VR-5 activity. Preferred proteins are VR-3  
15 proteins having at least one or more of the following domains: an ankyrin repeat domain, a pore domain, and/or a transmembrane domain, and, preferably, a VR-3 activity. Additional preferred VR-3 proteins have at least one ankyrin repeat domain and/or at least pore protein domain, and/or at least one transmembrane domain and are, preferably, encoded by a nucleic acid molecule having a nucleotide sequence which  
20 hybridizes under stringent hybridization conditions to a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1, 3, 4, or 6.

Accordingly, a further embodiment of the invention features isolated VR-5 polypeptides having a VR-5 activity. Preferred proteins are VR-5 proteins having at least one or more of the following domains: an ankyrin repeat domain, a pore domain, a  
25 transmembrane domain, and an ion transport protein domain, and, preferably, a VR-5 activity. Additional preferred proteins have at least one ankyrin repeat domain and/or at least one ion transport protein domain, and/or at least one pore domain, and/or at least one transmembrane domain, and are, preferably, encoded by a nucleic acid molecule having a nucleotide sequence which hybridizes under stringent hybridization conditions  
30 to a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1, 3, 4, or 6.

The nucleotide sequence of the isolated human VR-3 cDNA and the predicted amino acid sequence of the human VR-3 polypeptide are shown in Figure 1 and in SEQ ID NO:1 and SEQ ID NO:2, respectively.

The nucleotide sequence of the isolated human VR-5 cDNA and the predicted  
5 amino acid sequence of the human VR-5 polypeptide are shown in Figure 2 and in SEQ ID NO:4 and SEQ ID NO:5, respectively. A plasmid containing the nucleotide sequence encoding human VR-5 was deposited with the American Type Culture Collection (ATCC), 10801 University Boulevard, Manassas, VA 20110-2209, on June 8, 2000 and assigned Accession Number PTA-2013. This deposit will be maintained  
10 under the terms of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure. This deposits was made merely as a convenience for those of skill in the art and is not an admission that a deposit is required under 35 U.S.C. §112.

The human VR-3 gene, which is approximately 3026 nucleotides in length,  
15 encodes a protein having a molecular weight of approximately 79.8 kD and which is approximately 725 amino acid residues in length. The human VR-5 gene, which is approximately 3245 nucleotides in length, encodes a protein having a molecular weight of approximately 95.8 kD and which is approximately 871 amino acid residues in length.

20

Various aspects of the invention are described in further detail in the following subsections.

#### I. Isolated Nucleic Acid Molecules

25 One aspect of the invention pertains to isolated nucleic acid molecules that encode VR-3 or VR-5 proteins or biologically active portions thereof, as well as nucleic acid fragments sufficient for use as hybridization probes to identify VR-3- or VR-5-encoding nucleic acid molecules (*e.g.*, VR-3 mRNA, VR-5 mRNA) and fragments for use as PCR primers for the amplification or mutation of VR-3 or VR-5 nucleic acid  
30 molecules. As used herein, the term "nucleic acid molecule" is intended to include DNA molecules (*e.g.*, cDNA or genomic DNA) and RNA molecules (*e.g.*, mRNA) and analogs of the DNA or RNA generated using nucleotide analogs. The nucleic acid

molecule can be single-stranded or double-stranded, but preferably is double-stranded DNA.

The term "isolated nucleic acid molecule" includes nucleic acid molecules which are separated from other nucleic acid molecules which are present in the natural source of the nucleic acid. For example, with regards to genomic DNA, the term "isolated" includes nucleic acid molecules which are separated from the chromosome with which the genomic DNA is naturally associated. Preferably, an "isolated" nucleic acid is free of sequences which naturally flank the nucleic acid (*i.e.*, sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated VR-3 or VR-5 nucleic acid molecule can contain less than about 5 kb, 4kb, 3kb, 2kb, 1 kb, 0.5 kb or 0.1 kb of nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived. Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material, or culture medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized.

A nucleic acid molecule of the present invention, *e.g.*, a nucleic acid molecule having the nucleotide sequence of SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number PTA-2013, or a portion thereof, can be isolated using standard molecular biology techniques and the sequence information provided herein. Using all or portion of the nucleic acid sequence of SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number PTA-2013, as a hybridization probe, VR-3 or VR-5 nucleic acid molecules can be isolated using standard hybridization and cloning techniques (*e.g.*, as described in Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual*. 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

Moreover, a nucleic acid molecule encompassing all or a portion of SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number PTA-2013 can be isolated by the polymerase chain reaction (PCR) using synthetic oligonucleotide primers designed based upon the

sequence of SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number PTA-2013.

A nucleic acid of the invention can be amplified using cDNA, mRNA or alternatively, genomic DNA, as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis. Furthermore, oligonucleotides corresponding to VR-3 or VR-5 nucleotide sequences can be prepared by standard synthetic techniques, *e.g.*, using an automated DNA synthesizer.

10 In a preferred embodiment, an isolated nucleic acid molecule of the invention comprises the nucleotide sequence shown in SEQ ID NO:1. The sequence of SEQ ID NO:1 corresponds to the human VR-3 cDNA. This cDNA comprises sequences encoding the human VR-3 protein (*i.e.*, "the coding region", from nucleotides 280-2452), as well as 5' untranslated sequences (nucleotides 1-279) and 3' untranslated sequences (nucleotides 2453-3026). Alternatively, the nucleic acid molecule can  
15 comprise only the coding region of SEQ ID NO:1 (*e.g.*, nucleotides 1-2175, corresponding to SEQ ID NO:3).

In a further preferred embodiment, an isolated nucleic acid molecule of the invention comprises the nucleotide sequence shown in SEQ ID NO:4. The sequence of  
20 SEQ ID NO:4 corresponds to the human VR-5 cDNA. This cDNA comprises sequences encoding the human VR-5 protein (*i.e.*, "the coding region", from nucleotides 84-2696), as well as 5' untranslated sequences (nucleotides 1-83) and 3' untranslated sequences (nucleotides 2697-3245). Alternatively, the nucleic acid molecule can comprise only the coding region of SEQ ID NO:1 (*e.g.*, nucleotides 1-2613, corresponding to SEQ ID  
25 NO:6).

In another preferred embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule which is a complement of the nucleotide sequence shown in SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number PTA-2013, or a  
30 portion of any of these nucleotide sequences. A nucleic acid molecule which is complementary to the nucleotide sequence shown in SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as

Accession Number PTA-2013, is one which is sufficiently complementary to the nucleotide sequence shown in SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number PTA-2013, such that it can hybridize to the nucleotide sequence shown in SEQ ID NO:1, 3, 4, or 6,  
5 or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number PTA-2013, thereby forming a stable duplex.

In still another preferred embodiment, an isolated nucleic acid molecule of the present invention comprises a nucleotide sequence which is at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 83%, 85%, 86%, 90%, 91%, 92%, 95%, 96%, 97%, 98%,  
10 99%, 99.1%, 99.5% or more identical to the entire length of the nucleotide sequence shown in SEQ ID NO:1, 3, 4, or 6, or the entire length of the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number PTA-2013, or a portion of any of these nucleotide sequences.

Moreover, the nucleic acid molecule of the invention can comprise only a  
15 portion of the nucleic acid sequence of SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number PTA-2013, for example, a fragment which can be used as a probe or primer or a fragment encoding a portion of a VR-3 or VR-5 protein, *e.g.*, a biologically active portion of a VR-3 or VR-5 protein. The nucleotide sequence determined from the  
20 cloning of the VR-3 or VR-5 gene allows for the generation of probes and primers designed for use in identifying and/or cloning other VR-3 or VR-5 family members, as well as VR-3 or VR-5 homologues from other species. The probe/primer typically comprises substantially purified oligonucleotide. The oligonucleotide typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to  
25 at least about 12 or 15, preferably about 20 or 25, more preferably about 30, 35, 40, 45, 50, 55, 60, 65, or 75 consecutive nucleotides of a sense sequence of SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number PTA-2013, of an anti-sense sequence of SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as  
30 Accession Number PTA-2013, or of a naturally occurring allelic variant or mutant of SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number PTA-2013. In one embodiment, a nucleic

- 26 -

acid molecule of the present invention comprises a nucleotide sequence which is greater than 100-500, 500-1000, 1000-1500, 1500-2000, 2000-2500, 2500-3000, 3000-3200, or more nucleotides in length and hybridizes under stringent hybridization conditions to a nucleic acid molecule of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA  
5 insert of the plasmid deposited with ATCC as Accession Number PTA-2013.

Probes based on the VR-3 or VR-5 nucleotide sequences can be used to detect transcripts or genomic sequences encoding the same or homologous proteins. In preferred embodiments, the probe further comprises a label group attached thereto, *e.g.*, the label group can be a radioisotope, a fluorescent compound, an enzyme, or an enzyme  
10 co-factor. Such probes can be used as a part of a diagnostic test kit for identifying cells or tissue which misexpress a VR-3 or VR-5 protein, such as by measuring a level of a VR-3 or VR-5 -encoding nucleic acid in a sample of cells from a subject, *e.g.*, detecting VR-3 or VR-5 mRNA levels or determining whether a genomic VR-3 or VR-5 gene has been mutated or deleted.

15 A nucleic acid fragment encoding a "biologically active portion of a VR-3 or VR-5 protein" can be prepared by isolating a portion of the nucleotide sequence of SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number PTA-2013, which encodes a polypeptide having a VR-3 or VR-5 biological activity (the biological activities of the VR-3 or VR-5  
20 proteins are described herein), expressing the encoded portion of the VR-3 or VR-5 protein (*e.g.*, by recombinant expression *in vitro*) and assessing the activity of the encoded portion of the VR-3 or VR-5 protein.

The invention further encompasses nucleic acid molecules that differ from the nucleotide sequence shown in SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequence of  
25 the DNA insert of the plasmid deposited with ATCC as Accession Number PTA-2013, due to degeneracy of the genetic code and, thus, encode the same VR-3 or VR-5 proteins as those encoded by the nucleotide sequence shown in SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number PTA-2013. In another embodiment, an isolated nucleic acid  
30 molecule of the invention has a nucleotide sequence encoding a protein having an amino acid sequence shown in SEQ ID NO:2 or 5.

In addition to the VR-3 or VR-5 nucleotide sequences shown in SEQ ID NO:1, 3, 4, and 6, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number PTA-2013, it will be appreciated by those skilled in the art that DNA sequence polymorphisms that lead to changes in the amino acid sequences of the VR-3 or VR-5 proteins may exist within a population (*e.g.*, the human population). Such genetic polymorphism in the VR-3 or VR-5 genes may exist among individuals within a population due to natural allelic variation. As used herein, the terms "gene" and "recombinant gene" refer to nucleic acid molecules which include an open reading frame encoding a VR-3 or VR-5 protein, preferably a mammalian VR-3 or VR-5 protein, and can further include non-coding regulatory sequences, and introns.

Allelic variants of human VR-3 or VR-5 include both functional and non-functional VR-3 or VR-5 proteins. Functional allelic variants are naturally occurring amino acid sequence variants of the human VR-3 or VR-5 protein that maintain the ability to bind a VR-3 or VR-5 ligand or substrate and/or modulate pain signaling mechanisms, calcium homeostasis, cellular growth and/or proliferation, and/or tumorigenesis. Functional allelic variants will typically contain only conservative substitution of one or more amino acids of SEQ ID NO:2 or 5, or substitution, deletion or insertion of non-critical residues in non-critical regions of the protein.

Non-functional allelic variants are naturally occurring amino acid sequence variants of the human VR-3 or VR-5 proteins that do not have the ability to either bind a VR-3 or VR-5 ligand or substrate and/or modulate pain signaling mechanisms, calcium homeostasis mechanism, cellular growth and/or proliferation, and/or tumorigenesis. Non-functional allelic variants will typically contain a non-conservative substitution, a deletion, or insertion or premature truncation of the amino acid sequence of SEQ ID NO:2 or 5, or a substitution, insertion or deletion in critical residues or critical regions.

The present invention further provides non-human orthologues of the human VR-3 or VR-5 protein. Orthologues of the human VR-3 or VR-5 protein are proteins that are isolated from non-human organisms and possess the same VR-3 or VR-5 ligand binding and/or modulation of pain signaling mechanisms, modulation of calcium homeostasis mechanisms, modulation of cellular growth and/or proliferation, and/or modulation of tumorigenesis as the human VR-3 or VR-5 protein. Orthologues of the



human VR-3 or VR-5 protein can readily be identified as comprising an amino acid sequence that is substantially identical to SEQ ID NO:2 or 5.

Moreover, nucleic acid molecules encoding other VR-3 or VR-5 family members and, thus, which have a nucleotide sequence which differs from the VR-3 or VR-5 sequences of SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number PTA-2013 are intended to be within the scope of the invention. For example, another VR-3 or VR-5 cDNA can be identified based on the nucleotide sequence of human VR-3 or VR-5. Moreover, nucleic acid molecules encoding VR-3 or VR-5 proteins from different species, and which, thus, have a nucleotide sequence which differs from the VR-3 or VR-5 sequences of SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number PTA-2013 are intended to be within the scope of the invention. For example, a mouse VR-3 or VR-5 cDNA can be identified based on the nucleotide sequence of a human VR-3 or VR-5.

Nucleic acid molecules corresponding to natural allelic variants and homologues of the VR-3 or VR-5 cDNAs of the invention can be isolated based on their homology to the VR-3 or VR-5 nucleic acids disclosed herein using the cDNAs disclosed herein, or a portion thereof, as a hybridization probe according to standard hybridization techniques under stringent hybridization conditions. Nucleic acid molecules corresponding to natural allelic variants and homologues of the VR-3 or VR-5 cDNAs of the invention can further be isolated by mapping to the same chromosome or locus as the VR-3 or VR-5 gene.

Accordingly, in another embodiment, an isolated nucleic acid molecule of the invention is at least 15, 20, 25, 30 or more nucleotides in length and hybridizes under stringent conditions to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number PTA-2013. In other embodiment, the nucleic acid is at least 30, 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 1000, 1200, 1400, 1600, 1800, 2000, 2200, 2242, 2400, 2500, 2600, 2700, 2800, 2900, 3000, 3100, 3200 or more nucleotides in length.

As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences that are significantly identical or homologous to each other remain hybridized to each other. Preferably, the conditions are such that sequences at least about 70%, more preferably at least about 80%, even more preferably at least about 85% or 90% identical to each other remain hybridized to each other. Such stringent conditions are known to those skilled in the art and can be found in *Current Protocols in Molecular Biology*, Ausubel *et al.*, eds., John Wiley & Sons, Inc. (1995), sections 2, 4 and 6. Additional stringent conditions can be found in *Molecular Cloning: A Laboratory Manual*, Sambrook *et al.*, Cold Spring Harbor Press, Cold Spring Harbor, NY (1989), chapters 7, 9 and 11. A preferred, non-limiting example of stringent hybridization conditions includes hybridization in 4X sodium chloride/sodium citrate (SSC), at about 65-70°C (or hybridization in 4X SSC plus 50% formamide at about 42-50°C) followed by one or more washes in 1X SSC, at about 65-70°C. A preferred, non-limiting example of highly stringent hybridization conditions includes hybridization in 1X SSC, at about 65-70°C (or hybridization in 1X SSC plus 50% formamide at about 42-50°C) followed by one or more washes in 0.3X SSC, at about 65-70°C. A preferred, non-limiting example of reduced stringency hybridization conditions includes hybridization in 4X SSC, at about 50-60°C (or alternatively hybridization in 6X SSC plus 50% formamide at about 40-45°C) followed by one or more washes in 2X SSC, at about 50-60°C. Ranges intermediate to the above-recited values, *e.g.*, at 65-70°C or at 42-50°C are also intended to be encompassed by the present invention. SSPE (1xSSPE is 0.15M NaCl, 10mM NaH<sub>2</sub>PO<sub>4</sub>, and 1.25mM EDTA, pH 7.4) can be substituted for SSC (1xSSC is 0.15M NaCl and 15mM sodium citrate) in the hybridization and wash buffers; washes are performed for 15 minutes each after hybridization is complete. The hybridization temperature for hybrids anticipated to be less than 50 base pairs in length should be 5-10°C less than the melting temperature ( $T_m$ ) of the hybrid, where  $T_m$  is determined according to the following equations. For hybrids less than 18 base pairs in length,  $T_m(^{\circ}\text{C}) = 2(\# \text{ of A + T bases}) + 4(\# \text{ of G + C bases})$ . For hybrids between 18 and 49 base pairs in length,  $T_m(^{\circ}\text{C}) = 81.5 + 16.6(\log_{10}[\text{Na}^+]) + 0.41(\% \text{G+C}) - (600/\text{N})$ , where N is the number of bases in the hybrid, and  $[\text{Na}^+]$  is the concentration of sodium ions in the hybridization buffer ( $[\text{Na}^+]$  for 1xSSC = 0.165 M). It will also be recognized by the

skilled practitioner that additional reagents may be added to hybridization and/or wash buffers to decrease non-specific hybridization of nucleic acid molecules to membranes, for example, nitrocellulose or nylon membranes, including but not limited to blocking agents (*e.g.*, BSA or salmon or herring sperm carrier DNA), detergents (*e.g.*, SDS),  
5 chelating agents (*e.g.*, EDTA), Ficoll, PVP and the like. When using nylon membranes, in particular, an additional preferred, non-limiting example of stringent hybridization conditions is hybridization in 0.25-0.5M NaH<sub>2</sub>PO<sub>4</sub>, 7% SDS at about 65°C, followed by one or more washes at 0.02M NaH<sub>2</sub>PO<sub>4</sub>, 1% SDS at 65°C, see *e.g.*, Church and Gilbert (1984) *Proc. Natl. Acad. Sci. USA* 81:1991-1995, (or alternatively 0.2X SSC, 1% SDS).  
10 Preferably, an isolated nucleic acid molecule of the invention that hybridizes under stringent conditions to the sequence of SEQ ID NO:1, 3, 4, or 6 corresponds to a naturally-occurring nucleic acid molecule. As used herein, a "naturally-occurring" nucleic acid molecule refers to an RNA or DNA molecule having a nucleotide sequence that occurs in nature (*e.g.*, encodes a natural protein).

15 In addition to naturally-occurring allelic variants of the VR-3 or VR-5 sequences that may exist in the population, the skilled artisan will further appreciate that changes can be introduced by mutation into the nucleotide sequences of SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number PTA-2013, thereby leading to changes in the amino acid sequence of  
20 the encoded VR-3 or VR-5 proteins, without altering the functional ability of the VR-3 or VR-5 proteins. For example, nucleotide substitutions leading to amino acid substitutions at "non-essential" amino acid residues can be made in the sequence of SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number PTA-2013. A "non-essential" amino acid  
25 residue is a residue that can be altered from the wild-type sequence of VR-3 or VR-5 (*e.g.*, the sequence of SEQ ID NO:2 or 5) without altering the biological activity, whereas an "essential" amino acid residue is required for biological activity. For example, amino acid residues that are conserved among the VR-3 or VR-5 proteins of the present invention, *e.g.*, those present in the ankyrin repeat domain(s) or the ion  
30 transport protein domain(s) or the transmembrane domain(s), are predicted to be particularly unamenable to alteration. Furthermore, additional amino acid residues that

- 31 -

are conserved between the VR-3 or VR-5 proteins of the present invention and other members of the vanilloid receptor family are not likely to be amenable to alteration.

Accordingly, another aspect of the invention pertains to nucleic acid molecules encoding VR-3 or VR-5 proteins that contain changes in amino acid residues that are not essential for activity. Such VR-3 or VR-5 proteins differ in amino acid sequence from SEQ ID NO:2 or 5, yet retain biological activity. In one embodiment, the isolated nucleic acid molecule comprises a nucleotide sequence encoding a protein, wherein the protein comprises an amino acid sequence at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 86%, 90%, 91%, 92%, 95%, 96%, 97%, 97.1%, 98%, 99%, 99.1%, 99.5%, 99.9% or more identical to SEQ ID NO:2 or 5.

An isolated nucleic acid molecule encoding a VR-3 or VR-5 protein identical to the protein of SEQ ID NO:2 or 5, can be created by introducing one or more nucleotide substitutions, additions or deletions into the nucleotide sequence of SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number PTA-2013, such that one or more amino acid substitutions, additions or deletions are introduced into the encoded protein. Mutations can be introduced into SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number PTA-2013 by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis.

Preferably, conservative amino acid substitutions are made at one or more predicted non-essential amino acid residues. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (*e.g.*, lysine, arginine, histidine), acidic side chains (*e.g.*, aspartic acid, glutamic acid), uncharged polar side chains (*e.g.*, glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (*e.g.*, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (*e.g.*, threonine, valine, isoleucine) and aromatic side chains (*e.g.*, tyrosine, phenylalanine, tryptophan, histidine). Thus, a predicted nonessential amino acid residue in a VR-3 or VR-5 protein is preferably replaced with another amino acid residue from the same side chain family. Alternatively, in another embodiment, mutations can be introduced

randomly along all or part of a VR-3 or VR-5 coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for VR-3 or VR-5 biological activity to identify mutants that retain activity. Following mutagenesis of SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number PTA-2013, the encoded protein can be expressed recombinantly and the activity of the protein can be determined.

In another preferred embodiment, a mutant VR-3 or VR-5 protein can be assayed for the ability to (1) interact with a non-VR-3 or non-VR-5 protein molecule, *e.g.*, a VR-3 or VR-5 ligand or substrate; (2) activate a VR-3- or VR-5-dependent signal transduction pathway; (3) modulate calcium homeostasis mechanisms; (4) modulate membrane excitability; (5) modulate pain signaling mechanisms; (5) modulate cellular growth and/or proliferation; and (6) modulate tumorigenesis.

In addition to the nucleic acid molecules encoding VR-3 or VR-5 proteins described above, another aspect of the invention pertains to isolated nucleic acid molecules which are antisense thereto. An "antisense" nucleic acid comprises a nucleotide sequence which is complementary to a "sense" nucleic acid encoding a protein, *e.g.*, complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. Accordingly, an antisense nucleic acid can hydrogen bond to a sense nucleic acid. The antisense nucleic acid can be complementary to an entire VR-3 or VR-5 coding strand, or to only a portion thereof. In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence encoding VR-3 or VR-5. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues (*e.g.*, the coding region of human VR-3 or VR-5 corresponds to SEQ ID NO:3 or 6). In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence encoding VR-3 or VR-5. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (*i.e.*, also referred to as 5' and 3' untranslated regions).

Given the coding strand sequences encoding VR-3 or VR-5 disclosed herein (*e.g.*, SEQ ID NO:3 or 6), antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick base pairing. The antisense nucleic acid

molecule can be complementary to the entire coding region of VR-3 or VR-5 mRNA, but more preferably is an oligonucleotide which is antisense to only a portion of the coding or noncoding region of VR-3 or VR-5 mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of VR-3 or VR-5 mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis and enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (*e.g.*, an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, *e.g.*, phosphorothioate derivatives and acridine substituted nucleotides can be used. Examples of modified nucleotides which can be used to generate the antisense nucleic acid include 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xantine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (*v*), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (*v*), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)*w*, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (*i.e.*, RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

The antisense nucleic acid molecules of the invention are typically administered to a subject or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a VR-3 or VR-5 protein to thereby inhibit expression of the protein, *e.g.*, by inhibiting transcription and/or translation. The hybridization can be  
5 by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule which binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention include direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified  
10 to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, *e.g.*, by linking the antisense nucleic acid molecules to peptides or antibodies which bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the  
15 vectors described herein. To achieve sufficient intracellular concentrations of the antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an  $\alpha$ -anomeric nucleic acid molecule. An  $\alpha$ -anomeric nucleic acid molecule forms  
20 specific double-stranded hybrids with complementary RNA in which, contrary to the usual  $\beta$ -units, the strands run parallel to each other (Gaultier *et al.* (1987) *Nucleic Acids Res.* 15:6625-6641). The antisense nucleic acid molecule can also comprise a 2'-O-methylribonucleotide (Inoue *et al.* (1987) *Nucleic Acids Res.* 15:6131-6148) or a chimeric RNA-DNA analogue (Inoue *et al.* (1987) *FEBS Lett.* 215:327-330).

25 In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity which are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (*e.g.*, hammerhead ribozymes (described in Haselhoff and Gerlach (1988) *Nature* 334:585-591)) can be used to  
30 catalytically cleave VR-3 or VR-5 mRNA transcripts to thereby inhibit translation of VR-3 or VR-5 mRNA. A ribozyme having specificity for a VR-3- or VR-5-encoding nucleic acid can be designed based upon the nucleotide sequence of a VR-3 or VR-5

cDNA disclosed herein (*i.e.*, SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number PTA-2013). For example, a derivative of a *Tetrahymena* L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be  
5 cleaved in a VR-3- or VR-5-encoding mRNA. See, *e.g.*, Cech *et al.* U.S. Patent No. 4,987,071; and Cech *et al.* U.S. Patent No. 5,116,742. Alternatively, VR-3 or VR-5 mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, *e.g.*, Bartel, D. and Szostak, J.W. (1993) *Science* 261:1411-1418.

10 Alternatively, VR-3 or VR-5 gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory and/or 5' untranslated region of the VR-3 or VR-5 nucleotides (*e.g.*, the VR-3 or VR-5 promoter and/or enhancers; *e.g.*, nucleotides 1-277 of SEQ ID NO:1 or nucleotides 1-83 of SEQ ID NO:4) to form triple helical structures that prevent transcription of the VR-3 or VR-5 gene in target cells.  
15 See generally, Helene, C. (1991) *Anticancer Drug Des.* 6(6):569-84; Helene, C. *et al.* (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher, L.J. (1992) *Bioassays* 14(12):807-15.

In yet another embodiment, the VR-3 or VR-5 nucleic acid molecules of the present invention can be modified at the base moiety, sugar moiety or phosphate  
20 backbone to improve, *e.g.*, the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acid molecules can be modified to generate peptide nucleic acids (see Hyrup B. *et al.* (1996) *Bioorganic & Medicinal Chemistry* 4 (1): 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, *e.g.*, DNA mimics, in which the deoxyribose  
25 phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup B. *et al.* (1996) *supra*; Perry-O'Keefe *et al.*  
30 *Proc. Natl. Acad. Sci.* 93: 14670-675.



PNAs of VR-3 or VR-5 nucleic acid molecules can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, for example, inducing transcription or translation arrest or inhibiting replication. PNAs of VR-3 or VR-5 nucleic acid molecules can also be used in the analysis of single base pair mutations in a gene, (*e.g.*, by PNA-directed PCR clamping); as 'artificial restriction enzymes' when used in combination with other enzymes, (*e.g.*, S1 nucleases (Hyrup B. (1996) *supra*)); or as probes or primers for DNA sequencing or hybridization (Hyrup B. *et al.* (1996) *supra*; Perry-O'Keefe *supra*).

In another embodiment, PNAs of VR-3 or VR-5 can be modified, (*e.g.*, to enhance their stability or cellular uptake), by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras of VR-3 or VR-5 nucleic acid molecules can be generated which may combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, (*e.g.*, RNase H and DNA polymerases), to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup B. (1996) *supra*). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup B. (1996) *supra* and Finn P.J. *et al.* (1996) *Nucleic Acids Res.* 24 (17): 3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry and modified nucleoside analogs, *e.g.*, 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used as a between the PNA and the 5' end of DNA (Mag, M. *et al.* (1989) *Nucleic Acid Res.* 17: 5973-88). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn P.J. *et al.* (1996) *supra*). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment (Peterser, K.H. *et al.* (1975) *Bioorganic Med. Chem. Lett.* 5: 1119-11124).

In other embodiments, the oligonucleotide may include other appended groups such as peptides (*e.g.*, for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see, *e.g.*, Letsinger *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86:6553-6556; Lemaitre *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:648-652; 5 PCT Publication No. W088/09810) or the blood-brain barrier (see, *e.g.*, PCT Publication No. W089/10134). In addition, oligonucleotides can be modified with hybridization-triggered cleavage agents (See, *e.g.*, Krol *et al.* (1988) *Bio-Techniques* 6:958-976) or intercalating agents (See, *e.g.*, Zon (1988) *Pharm. Res.* 5:539-549). To this end, the oligonucleotide may be conjugated to another molecule, (*e.g.*, a peptide, hybridization 10 triggered cross-linking agent, transport agent, or hybridization-triggered cleavage agent).

## II. Isolated VR-3 or VR-5 Proteins and Anti-VR-3 or Anti-VR-5 Antibodies

One aspect of the invention pertains to isolated VR-3 or VR-5 proteins, and biologically active portions thereof, as well as polypeptide fragments suitable for use as 15 immunogens to raise anti-VR-3 or anti-VR-5 antibodies. In one embodiment, native VR-3 or VR-5 proteins can be isolated from cells or tissue sources by an appropriate purification scheme using standard protein purification techniques. In another embodiment, VR-3 or VR-5 proteins are produced by recombinant DNA techniques. Alternative to recombinant expression, a VR-3 or VR-5 protein or polypeptide can be 20 synthesized chemically using standard peptide synthesis techniques.

An "isolated" or "purified" protein or biologically active portion thereof is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the VR-3 or VR-5 protein is derived, or substantially free from chemical precursors or other chemicals when chemically synthesized. The language 25 "substantially free of cellular material" includes preparations of VR-3 or VR-5 protein in which the protein is separated from cellular components of the cells from which it is isolated or recombinantly produced. In one embodiment, the language "substantially free of cellular material" includes preparations of VR-3 or VR-5 protein having less than about 30% (by dry weight) of non-VR-3 or non-VR-5 protein (also referred to herein as 30 a "contaminating protein"), more preferably less than about 20% of non-VR-3 or non-VR-5 protein, still more preferably less than about 10% of non-VR-3 or non-VR-5 protein, and most preferably less than about 5% non-VR-3 or non-VR-5 protein. When

the VR-3 or VR-5 protein or biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, *i.e.*, culture medium represents less than about 20%, more preferably less than about 10%, and most preferably less than about 5% of the volume of the protein preparation.

5           The language "substantially free of chemical precursors or other chemicals" includes preparations of VR-3 or VR-5 protein in which the protein is separated from chemical precursors or other chemicals which are involved in the synthesis of the protein. In one embodiment, the language "substantially free of chemical precursors or other chemicals" includes preparations of VR-3 or VR-5 protein having less than about  
10   30% (by dry weight) of chemical precursors or non-VR-3 or non-VR-5 chemicals, more preferably less than about 20% chemical precursors or non-VR-3 or non-VR-5 chemicals, still more preferably less than about 10% chemical precursors or non-VR-3 or non-VR-5 chemicals, and most preferably less than about 5% chemical precursors or non-VR-3 or non-VR-5 chemicals.

15           As used herein, a "biologically active portion" of a VR-3 or VR-5 protein includes a fragment of a VR-3 or VR-5 protein which participates in an interaction between a VR-3 or VR-5 molecule and a non-VR-3 or non-VR-5 molecule. Biologically active portions of a VR-3 or VR-5 protein include peptides comprising amino acid sequences sufficiently identical to or derived from the amino acid sequence  
20   of the VR-3 or VR-5 protein, *e.g.*, the amino acid sequence shown in SEQ ID NO:2 or 5, which include less amino acids than the full length VR-3 or VR-5 proteins, and exhibit at least one activity of a VR-3 or VR-5 protein. Typically, biologically active portions comprise a domain or motif with at least one activity of the VR-3 or VR-5 protein, *e.g.*, modulating pain signaling mechanisms, modulating calcium homeostasis, and/or  
25   modulating cellular growth and/or proliferation, or modulation of tumorigenesis. A biologically active portion of a VR-3 or VR-5 protein can be a polypeptide which is, for example, 10, 25, 50, 100, 200, 500, or more amino acids in length. Biologically active portions of a VR-3 or VR-5 protein can be used as targets for developing agents which modulate a VR-3 or VR-5 mediated activity, *e.g.*, a pain signaling mechanism, a calcium  
30   homeostasis mechanism, cellular growth and/or proliferation, or tumorigenesis.

the VR-3 or VR-5 protein or biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, *i.e.*, culture medium represents less than about 20%, more preferably less than about 10%, and most preferably less than about 5% of the volume of the protein preparation.

5           The language "substantially free of chemical precursors or other chemicals" includes preparations of VR-3 or VR-5 protein in which the protein is separated from chemical precursors or other chemicals which are involved in the synthesis of the protein. In one embodiment, the language "substantially free of chemical precursors or other chemicals" includes preparations of VR-3 or VR-5 protein having less than about  
10   30% (by dry weight) of chemical precursors or non-VR-3 or non-VR-5 chemicals, more preferably less than about 20% chemical precursors or non-VR-3 or non-VR-5 chemicals, still more preferably less than about 10% chemical precursors or non-VR-3 or non-VR-5 chemicals, and most preferably less than about 5% chemical precursors or non-VR-3 or non-VR-5 chemicals.

15           As used herein, a "biologically active portion" of a VR-3 or VR-5 protein includes a fragment of a VR-3 or VR-5 protein which participates in an interaction between a VR-3 or VR-5 molecule and a non-VR-3 or non-VR-5 molecule. Biologically active portions of a VR-3 or VR-5 protein include peptides comprising amino acid sequences sufficiently identical to or derived from the amino acid sequence  
20   of the VR-3 or VR-5 protein, *e.g.*, the amino acid sequence shown in SEQ ID NO:2 or 5, which include less amino acids than the full length VR-3 or VR-5 proteins, and exhibit at least one activity of a VR-3 or VR-5 protein. Typically, biologically active portions comprise a domain or motif with at least one activity of the VR-3 or VR-5 protein, *e.g.*,  
25   modulating pain signaling mechanisms, modulating calcium homeostasis, and/or modulating cellular growth and/or proliferation, or modulation of tumorigenesis. A biologically active portion of a VR-3 or VR-5 protein can be a polypeptide which is, for example, 10, 25, 50, 100, 200, 500, or more amino acids in length. Biologically active portions of a VR-3 or VR-5 protein can be used as targets for developing agents which  
30   modulate a VR-3 or VR-5 mediated activity, *e.g.*, a pain signaling mechanism, a calcium homeostasis mechanism, cellular growth and/or proliferation, or tumorigenesis.

In one embodiment, a biologically active portion of a VR-3 or VR-5 protein comprises at least one ankyrin repeat domain, and/or at least one transmembrane domain and/or at least one ion transport protein domain. It is to be understood that a preferred biologically active portion of a VR-3 or VR-5 protein of the present invention may  
5 contain at least one ankyrin repeat domain. Another preferred biologically active portion of a VR-5 protein may contain a ion transport protein domain. Another preferred biologically active portion of a VR-3 or VR-5 protein may contain at least one transmembrane domain. Moreover, other biologically active portions, in which other regions of the protein are deleted, can be prepared by recombinant techniques and  
10 evaluated for one or more of the functional activities of a native VR-3 or VR-5 protein.

In a preferred embodiment, the VR-3 or VR-5 protein has an amino acid sequence shown in SEQ ID NO:2 or 5. In other embodiments, the VR-3 or VR-5 protein is substantially identical to SEQ ID NO:2 or 5, and retains the functional activity of the protein of SEQ ID NO:2 or 5, yet differs in amino acid sequence due to natural  
15 allelic variation or mutagenesis, as described in detail in subsection I above. Accordingly, in another embodiment, the VR-3 or VR-5 protein is a protein which comprises an amino acid sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 86%, 90%, 91%, 92%, 95%, 96%, 97%, 97.1%, 98%, 99%, 99.1%, 99.5%, 99.9% or more identical to SEQ ID NO:2 or 5.

20 To determine the percent identity of two amino acid sequences or of two nucleic acid sequences, the sequences are aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in one or both of a first and a second amino acid or nucleic acid sequence for optimal alignment and non-identical sequences can be disregarded for comparison purposes). In a preferred embodiment, the length of a reference sequence  
25 aligned for comparison purposes is at least 30%, preferably at least 40%, more preferably at least 50%, even more preferably at least 60%, and even more preferably at least 70%, 80%, or 90% of the length of the reference sequence (*e.g.*, when aligning a second sequence to the VR-3 amino acid sequence of SEQ ID NO:2 having 725 amino acid residues, at least 218, preferably at least 310, more preferably at least 388, even  
30 more preferably at least 465, and even more preferably at least 543, 620 or 698 amino acid residues are aligned). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the

first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position (as used herein amino acid or nucleic acid "identity" is equivalent to amino acid or nucleic acid "homology"). The percent identity between the two sequences is a  
5 function of the number of identical positions shared by the sequences, taking into account the number of gaps, and the length of each gap, which need to be introduced for optimal alignment of the two sequences.

The comparison of sequences and determination of percent identity between two sequences can be accomplished using a mathematical algorithm. In a preferred  
10 embodiment, the percent identity between two amino acid sequences is determined using the Needleman and Wunsch (*J. Mol. Biol.* (48):444-453 (1970)) algorithm which has been incorporated into the GAP program in the GCG software package (available at <http://www.gcg.com>), using either a Blosum 62 matrix or a PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6, or 4 and a length weight of 1, 2, 3, 4, 5, or 6. In yet  
15 another preferred embodiment, the percent identity between two nucleotide sequences is determined using the GAP program in the GCG software package (available at <http://www.gcg.com>), using a NWSgapdna.CMP matrix and a gap weight of 40, 50, 60, 70, or 80 and a length weight of 1, 2, 3, 4, 5, or 6. In another embodiment, the percent identity between two amino acid or nucleotide sequences is determined using the  
20 algorithm of E. Meyers and W. Miller (Myers and Miller, *Comput. Appl. Biosci.* 4:11-17 (1988)) which has been incorporated into the ALIGN program (version 2.0), using a PAM120 weight residue table, a gap length penalty of 12 and a gap penalty of 4.

The nucleic acid and protein sequences of the present invention can further be used as a "query sequence" to perform a search against public databases to, for example,  
25 identify other family members or related sequences. Such searches can be performed using the NBLAST and XBLAST programs (version 2.0) of Altschul, *et al.* (1990) *J. Mol. Biol.* 215:403-10. BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to VR-3 or VR-5 nucleic acid molecules of the invention. BLAST protein  
30 searches can be performed with the XBLAST program, score = 100, wordlength = 3 to obtain amino acid sequences homologous to VR-3 or VR-5 protein molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can

be utilized as described in Altschul *et al.*, (1997) *Nucleic Acids Res.* 25(17):3389-3402. When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (*e.g.*, XBLAST and NBLAST) can be used. See <http://www.ncbi.nlm.nih.gov>.

5           The invention also provides VR-3 or VR-5 chimeric or fusion proteins. As used herein, a VR-3 or VR-5 "chimeric protein" or "fusion protein" comprises a VR-3 or VR-5 polypeptide operatively linked to a non-VR-3 or non-VR-5 polypeptide. A "VR-3 polypeptide" or a "VR-5 polypeptide" includes a polypeptide having an amino acid sequence corresponding to VR-3 or VR-5, whereas a "non-VR-3 peptide" or a "non-VR-5 polypeptide" includes a polypeptide having an amino acid sequence corresponding to  
10 a protein which is not substantially homologous to a VR-3 or VR-5 protein, *e.g.*, a protein which is different from the VR-3 or VR-5 protein and which is derived from the same or a different organism. Within a VR-3 or VR-5 fusion protein the VR-3 or VR-5 polypeptide can correspond to all or a portion of a VR-3 or VR-5 protein. In a preferred  
15 embodiment, a VR-3 or VR-5 fusion protein comprises at least one biologically active portion of a VR-3 or VR-5 protein. In another preferred embodiment, a VR-3 or VR-5 fusion protein comprises at least two biologically active portions of a VR-3 or VR-5 protein. Within the fusion protein, the term "operatively linked" is intended to indicate that the VR-3 or VR-5 polypeptide and the non-VR-3 or non-VR-5 polypeptide are  
20 fused in-frame to each other. The non-VR-3 or non-VR-5 polypeptide can be fused to the N-terminus or C-terminus of the VR-3 or VR-5 polypeptide.

For example, in one embodiment, the fusion protein is a GST-VR-3 or GST-VR-5 fusion protein in which the VR-3 or VR-5 sequences are fused to the C-terminus of the GST sequences. Such fusion proteins can facilitate the purification of  
25 recombinant VR-3 or VR-5.

In another embodiment, the fusion protein is a VR-3 or VR-5 protein containing a heterologous signal sequence at its N-terminus. In certain host cells (*e.g.*, mammalian host cells), expression and/or secretion of VR-3 or VR-5 can be increased through use of a heterologous signal sequence.

30           The VR-3 or VR-5 fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject *in vivo*. The VR-3 or VR-5 fusion proteins can be used to affect the bioavailability of a VR-3 or VR-5 ligand or

substrate. Use of VR-3 or VR-5 fusion proteins may be useful therapeutically for the treatment of disorders caused by, for example, (i) aberrant modification or mutation of a gene encoding a VR-3 or VR-5 protein; (ii) mis-regulation of the VR-3 or VR-5 gene; and (iii) aberrant post-translational modification of a VR-3 or VR-5 protein.

5           Moreover, the VR-3-or VR-5-fusion proteins of the invention can be used as immunogens to produce anti-VR-3 or anti-VR-5 antibodies in a subject, to purify VR-3 or VR-5 ligands and in screening assays to identify molecules which inhibit the interaction of VR-3 or VR-5 with a VR-3 or VR-5 ligand or substrate.

          Preferably, a VR-3 or VR-5 chimeric or fusion protein of the invention is  
10   produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, for example by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase  
15   treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene fragments which can subsequently be annealed and reamplified to  
20   generate a chimeric gene sequence (see, for example, *Current Protocols in Molecular Biology*, eds. Ausubel *et al.* John Wiley & Sons: 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (*e.g.*, a GST polypeptide). A VR-3- or VR-5-encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the VR-3 or VR-5  
25   protein.

          The present invention also pertains to variants of the VR-3 or VR-5 proteins which function as either VR-3 or VR-5 agonists (mimetics) or as VR-3 or VR-5 antagonists. Variants of the VR-3 or VR-5 proteins can be generated by mutagenesis, *e.g.*, discrete point mutation or truncation of a VR-3 or VR-5 protein. An agonist of the  
30   VR-3 or VR-5 proteins can retain substantially the same, or a subset, of the biological activities of the naturally occurring form of a VR-3 or VR-5 protein. An antagonist of a VR-3 or VR-5 protein can inhibit one or more of the activities of the naturally occurring



form of the VR-3 or VR-5 protein by, for example, competitively modulating a VR-3- or VR-5-mediated activity of a VR-3 or VR-5 protein. Thus, specific biological effects can be elicited by treatment with a variant of limited function. In one embodiment, treatment of a subject with a variant having a subset of the biological activities of the naturally occurring form of the protein has fewer side effects in a subject relative to treatment with the naturally occurring form of the VR-3 or VR-5 protein.

In one embodiment, variants of a VR-3 or VR-5 protein which function as either VR-3 or VR-5 agonists (mimetics) or as VR-3 or VR-5 antagonists can be identified by screening combinatorial libraries of mutants, *e.g.*, truncation mutants, of a VR-3 or VR-5 protein for VR-3 or VR-5 protein agonist or antagonist activity. In one embodiment, a variegated library of VR-3 or VR-5 variants is generated by combinatorial mutagenesis at the nucleic acid level and is encoded by a variegated gene library. A variegated library of VR-3 or VR-5 variants can be produced by, for example, enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of potential VR-3 or VR-5 sequences is expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (*e.g.*, for phage display) containing the set of VR-3 or VR-5 sequences therein. There are a variety of methods which can be used to produce libraries of potential VR-3 or VR-5 variants from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be performed in an automatic DNA synthesizer, and the synthetic gene then ligated into an appropriate expression vector. Use of a degenerate set of genes allows for the provision, in one mixture, of all of the sequences encoding the desired set of potential VR-3 or VR-5 sequences. Methods for synthesizing degenerate oligonucleotides are known in the art (see, *e.g.*, Narang, S.A. (1983) *Tetrahedron* 39:3; Itakura *et al.* (1984) *Annu. Rev. Biochem.* 53:323; Itakura *et al.* (1984) *Science* 198:1056; Ike *et al.* (1983) *Nucleic Acid Res.* 11:477).

In addition, libraries of fragments of a VR-3 or VR-5 protein coding sequence can be used to generate a variegated population of VR-3 or VR-5 fragments for screening and subsequent selection of variants of a VR-3 or VR-5 protein. In one embodiment, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of a VR-3 or VR-5 coding sequence with a nuclease under conditions wherein nicking occurs only about once per molecule, denaturing the

double stranded DNA, renaturing the DNA to form double stranded DNA which can include sense/antisense pairs from different nicked products, removing single stranded portions from reformed duplexes by treatment with S1 nuclease, and ligating the resulting fragment library into an expression vector. By this method, an expression  
5 library can be derived which encodes N-terminal, C-terminal and internal fragments of various sizes of the VR-3 or VR-5 protein.

Several techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products having a selected property. Such techniques are adaptable for  
10 rapid screening of the gene libraries generated by the combinatorial mutagenesis of VR-3 or VR-5 proteins. The most widely used techniques, which are amenable to high through-put analysis, for screening large gene libraries typically include cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in  
15 which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recursive ensemble mutagenesis (REM), a new technique which enhances the frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify VR-3 or VR-5 variants (Arkin and Yourvan (1992) *Proc. Natl. Acad. Sci. USA* 89:7811-7815; Delgrave *et al.* (1993)  
20 *Protein Engineering* 6(3):327-331).

In one embodiment, cell based assays can be exploited to analyze a variegated VR-3 or VR-5 library. For example, a library of expression vectors can be transfected into a cell line, *e.g.*, a neuronal cell line, which ordinarily responds to VR-3 or VR-5 in a particular VR-3 or VR-5 ligand-dependent manner. The transfected cells are then  
25 contacted with a VR-3 or VR-5 ligand and the effect of expression of the mutant on signaling by the VR-3 or VR-5 ligand can be detected, *e.g.*, by monitoring intracellular calcium concentration, neuronal membrane depolarization, or the activity of a VR-3- or VR5-regulated transcription factor. Plasmid DNA can then be recovered from the cells which score for inhibition, or alternatively, potentiation of signaling by the VR-3 or VR-  
30 5 ligand, and the individual clones further characterized. In related cell-based assays, changes in membrane potential can be measured in live cells which express VR-3 or VR-5 molecules of the invention. Such an assay can be used for screening compound

libraries for useful ligands which interact with VR-3 or VR-5, or can be used to identify variants of VR-3 and VR-5 which have useful properties. Other cell based assay include those which can monitor fluxes in intracellular calcium levels, *e.g.*, flow cytometry (Valet and Raffael, 1985, *Naturwiss.*, 72:600-602). Also within the scope of the invention are assays and models which utilize VR-3 or VR-5 nucleic acids to create transgenic organisms for identifying useful pharmaceutical compounds or variants of the VR-3 and/or VR-5 molecules.

An isolated VR-3 or VR-5 protein, or a portion or fragment thereof, can be used as an immunogen to generate antibodies that bind VR-3 or VR-5 using standard techniques for polyclonal and monoclonal antibody preparation. A full-length VR-3 or VR-5 protein can be used or, alternatively, the invention provides antigenic peptide fragments of VR-3 or VR-5 for use as immunogens. The antigenic peptide of VR-3 or VR-5 comprises at least 8 amino acid residues of the amino acid sequence shown in SEQ ID NO:2 or 5 and encompasses an epitope of VR-3 or VR-5 such that an antibody raised against the peptide forms a specific immune complex with VR-3 or VR-5. Preferably, the antigenic peptide comprises at least 10 amino acid residues, more preferably at least 15 amino acid residues, even more preferably at least 20 amino acid residues, and most preferably at least 30 amino acid residues.

Preferred epitopes encompassed by the antigenic peptide are regions of VR-3 or VR-5 that are located on the surface of the protein, *e.g.*, hydrophilic regions, as well as regions with high antigenicity (see, for example, Figures 6 and 7).

A VR-3 or VR-5 immunogen typically is used to prepare antibodies by immunizing a suitable subject, (*e.g.*, rabbit, goat, mouse or other mammal) with the immunogen. An appropriate immunogenic preparation can contain, for example, recombinantly expressed VR-3 or VR-5 protein or a chemically synthesized VR-3 or VR-5 polypeptide. The preparation can further include an adjuvant, such as Freund's complete or incomplete adjuvant, or similar immunostimulatory agent. Immunization of a suitable subject with an immunogenic VR-3 or VR-5 preparation induces a polyclonal anti-VR-3 or anti-VR-5 antibody response.

Accordingly, another aspect of the invention pertains to anti-VR-3 or anti-VR-5 antibodies. The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin molecules, *i.e.*, molecules that

contain an antigen binding site which specifically binds (immunoreacts with) an antigen, such as VR-3 or VR-5. Examples of immunologically active portions of immunoglobulin molecules include F(ab) and F(ab')<sub>2</sub> fragments which can be generated by treating the antibody with an enzyme such as pepsin. The invention provides

5 polyclonal and monoclonal antibodies that bind VR-3 or VR-5. The term "monoclonal antibody" or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one species of an antigen binding site capable of immunoreacting with a particular epitope of VR-3 or VR-5. A monoclonal antibody composition thus typically displays a single binding affinity for a particular VR-3 or

10 VR-5 protein with which it immunoreacts.

Polyclonal anti-VR-3 or anti-VR-5 antibodies can be prepared as described above by immunizing a suitable subject with a VR-3 or VR-5 immunogen. The anti-VR-3 or anti-VR-5 antibody titer in the immunized subject can be monitored over time by standard techniques, such as with an enzyme linked immunosorbent assay (ELISA)

15 using immobilized VR-3 or VR-5. If desired, the antibody molecules directed against VR-3 or VR-5 can be isolated from the mammal (*e.g.*, from the blood) and further purified by well known techniques, such as protein A chromatography to obtain the IgG fraction. At an appropriate time after immunization, *e.g.*, when the anti-VR-3 or anti-VR-5 antibody titers are highest, antibody-producing cells can be obtained from the

20 subject and used to prepare monoclonal antibodies by standard techniques, such as the hybridoma technique originally described by Kohler and Milstein (1975) *Nature* 256:495-497) (see also, Brown *et al.* (1981) *J. Immunol.* 127:539-46; Brown *et al.* (1980) *J. Biol. Chem.* 255:4980-83; Yeh *et al.* (1976) *Proc. Natl. Acad. Sci. USA* 76:2927-31; and Yeh *et al.* (1982) *Int. J. Cancer* 29:269-75), the more recent human B

25 cell hybridoma technique (Kozbor *et al.* (1983) *Immunol Today* 4:72), the EBV-hybridoma technique (Cole *et al.* (1985), *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc., pp. 77-96) or trioma techniques. The technology for producing monoclonal antibody hybridomas is well known (see generally R. H. Kenneth, in *Monoclonal Antibodies: A New Dimension In Biological Analyses*, Plenum Publishing

30 Corp., New York, New York (1980); E. A. Lerner (1981) *Yale J. Biol. Med.*, 54:387-402; M. L. Gefter *et al.* (1977) *Somatic Cell Genet.* 3:231-36). Briefly, an immortal cell line (typically a myeloma) is fused to lymphocytes (typically splenocytes)

from a mammal immunized with a VR-3 or VR-5 immunogen as described above, and the culture supernatants of the resulting hybridoma cells are screened to identify a hybridoma producing a monoclonal antibody that binds VR-3 or VR-5.

Any of the many well known protocols used for fusing lymphocytes and  
5 immortalized cell lines can be applied for the purpose of generating an anti-VR-3 or anti-VR-5 monoclonal antibody (see, *e.g.*, G. Galfre *et al.* (1977) *Nature* 266:55052; Gefter *et al.* *Somatic Cell Genet.*, cited *supra*; Lerner, *Yale J. Biol. Med.*, cited *supra*; Kenneth, *Monoclonal Antibodies*, cited *supra*). Moreover, the ordinarily skilled worker will appreciate that there are many variations of such methods which also would be  
10 useful. Typically, the immortal cell line (*e.g.*, a myeloma cell line) is derived from the same mammalian species as the lymphocytes. For example, murine hybridomas can be made by fusing lymphocytes from a mouse immunized with an immunogenic preparation of the present invention with an immortalized mouse cell line. Preferred immortal cell lines are mouse myeloma cell lines that are sensitive to culture medium  
15 containing hypoxanthine, aminopterin and thymidine ("HAT medium"). Any of a number of myeloma cell lines can be used as a fusion partner according to standard techniques, *e.g.*, the P3-NS1/1-Ag4-1, P3-x63-Ag8.653 or Sp2/O-Ag14 myeloma lines. These myeloma lines are available from ATCC. Typically, HAT-sensitive mouse myeloma cells are fused to mouse splenocytes using polyethylene glycol ("PEG").  
20 Hybridoma cells resulting from the fusion are then selected using HAT medium, which kills unfused and unproductively fused myeloma cells (unfused splenocytes die after several days because they are not transformed). Hybridoma cells producing a monoclonal antibody of the invention are detected by screening the hybridoma culture supernatants for antibodies that bind VR-3 or VR-5, *e.g.*, using a standard ELISA assay.

25 Alternative to preparing monoclonal antibody-secreting hybridomas, a monoclonal anti-VR-3 or anti-VR-5 antibody can be identified and isolated by screening a recombinant combinatorial immunoglobulin library (*e.g.*, an antibody phage display library) with VR-3 or VR-5 to thereby isolate immunoglobulin library members that bind VR-3 or VR-5. Kits for generating and screening phage display libraries are  
30 commercially available (*e.g.*, the Pharmacia *Recombinant Phage Antibody System*, Catalog No. 27-9400-01; and the Stratagene *SurfZAP™ Phage Display Kit*, Catalog No. 240612). Additionally, examples of methods and reagents particularly amenable for use

in generating and screening antibody display library can be found in, for example, Ladner *et al.* U.S. Patent No. 5,223,409; Kang *et al.* PCT International Publication No. WO 92/18619; Dower *et al.* PCT International Publication No. WO 91/17271; Winter *et al.* PCT International Publication WO 92/20791; Markland *et al.* PCT International Publication No. WO 92/15679; Breitling *et al.* PCT International Publication WO 93/01288; McCafferty *et al.* PCT International Publication No. WO 92/01047; Garrard *et al.* PCT International Publication No. WO 92/09690; Ladner *et al.* PCT International Publication No. WO 90/02809; Fuchs *et al.* (1991) *Bio/Technology* 9:1370-1372; Hay *et al.* (1992) *Hum. Antibod. Hybridomas* 3:81-85; Huse *et al.* (1989) *Science* 246:1275-1281; Griffiths *et al.* (1993) *EMBO J* 12:725-734; Hawkins *et al.* (1992) *J. Mol. Biol.* 226:889-896; Clarkson *et al.* (1991) *Nature* 352:624-628; Gram *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:3576-3580; Garrad *et al.* (1991) *Bio/Technology* 9:1373-1377; Hoogenboom *et al.* (1991) *Nuc. Acid Res.* 19:4133-4137; Barbas *et al.* (1991) *Proc. Natl. Acad. Sci. USA* 88:7978-7982; and McCafferty *et al.* *Nature* (1990) 348:552-554.

Additionally, recombinant anti-VR-3 or anti-VR-5 antibodies, such as chimeric and humanized monoclonal antibodies, comprising both human and non-human portions, which can be made using standard recombinant DNA techniques, are within the scope of the invention. Such chimeric and humanized monoclonal antibodies can be produced by recombinant DNA techniques known in the art, for example using methods described in Robinson *et al.* International Application No. PCT/US86/02269; Akira, *et al.* European Patent Application 184,187; Taniguchi, M., European Patent Application 171,496; Morrison *et al.* European Patent Application 173,494; Neuberger *et al.* PCT International Publication No. WO 86/01533; Cabilly *et al.* U.S. Patent No. 4,816,567; Cabilly *et al.* European Patent Application 125,023; Better *et al.* (1988) *Science* 240:1041-1043; Liu *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:3439-3443; Liu *et al.* (1987) *J. Immunol.* 139:3521-3526; Sun *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:214-218; Nishimura *et al.* (1987) *Canc. Res.* 47:999-1005; Wood *et al.* (1985) *Nature* 314:446-449; and Shaw *et al.* (1988) *J. Natl. Cancer Inst.* 80:1553-1559; Morrison, S. L. (1985) *Science* 229:1202-1207; Oi *et al.* (1986) *BioTechniques* 4:214; Winter U.S. Patent 5,225,539; Jones *et al.* (1986) *Nature* 321:552-525; Verhoeven *et al.* (1988) *Science* 239:1534; and Beidler *et al.* (1988) *J. Immunol.* 141:4053-4060.

An anti-VR-3 or anti-VR-5 antibody (*e.g.*, monoclonal antibody) can be used to isolate VR-3 or VR-5 by standard techniques, such as affinity chromatography or immunoprecipitation. An anti-VR-3 or anti-VR-5 antibody can facilitate the purification of natural VR-3 or VR-5 from cells and of recombinantly produced VR-3 or VR-5 expressed in host cells. Moreover, an anti-VR-3 or anti-VR-5 antibody can be used to detect VR-3 or VR-5 protein (*e.g.*, in a cellular lysate or cell supernatant) in order to evaluate the abundance and pattern of expression of the VR-3 or VR-5 protein. Anti-VR-3 or anti-VR-5 antibodies can be used diagnostically to monitor protein levels in tissue as part of a clinical testing procedure, *e.g.*, to, for example, determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling (*i.e.*, physically linking) the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase,  $\beta$ -galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include  $^{125}\text{I}$ ,  $^{131}\text{I}$ ,  $^{35}\text{S}$  or  $^3\text{H}$ .

### III. Recombinant Expression Vectors and Host Cells

Another aspect of the invention pertains to vectors, preferably expression vectors, containing a nucleic acid encoding a VR-3 or VR-5 protein (or a portion thereof). As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (*e.g.*, bacterial vectors having a bacterial origin of replication and episomal mammalian

vectors). Other vectors (*e.g.*, non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to  
5 herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and "vector" can be used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors, such as viral vectors (*e.g.*, replication defective  
10 retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

The recombinant expression vectors of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell, which means that the recombinant expression vectors include one or more regulatory  
15 sequences, selected on the basis of the host cells to be used for expression, which is operatively linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operably linked" is intended to mean that the nucleotide sequence of interest is linked to the regulatory sequence(s) in a manner which allows for expression of the nucleotide sequence (*e.g.*, in an *in vitro* transcription/translation system or in a  
20 host cell when the vector is introduced into the host cell). The term "regulatory sequence" is intended to include promoters, enhancers and other expression control elements (*e.g.*, polyadenylation signals). Such regulatory sequences are described, for example, in Goeddel; *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Regulatory sequences include those which  
25 direct constitutive expression of a nucleotide sequence in many types of host cells and those which direct expression of the nucleotide sequence only in certain host cells (*e.g.*, tissue-specific regulatory sequences). It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, and the like. The  
30 expression vectors of the invention can be introduced into host cells to thereby produce proteins or peptides, including fusion proteins or peptides, encoded by nucleic acids as



described herein (*e.g.*, VR-3 or VR-5 proteins, mutant forms of VR-3 or VR-5 proteins, fusion proteins, and the like).

The recombinant expression vectors of the invention can be designed for expression of VR-3 or VR-5 proteins in prokaryotic or eukaryotic cells. For example, VR-3 or VR-5 proteins can be expressed in bacterial cells such as *E. coli*, insect cells (using baculovirus expression vectors) yeast cells or mammalian cells. Suitable host cells are discussed further in Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

Expression of proteins in prokaryotes is most often carried out in *E. coli* with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, usually to the amino terminus of the recombinant protein. Such fusion vectors typically serve three purposes: 1) to increase expression of recombinant protein; 2) to increase the solubility of the recombinant protein; and 3) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith, D.B. and Johnson, K.S. (1988) *Gene* 67:31-40), pMAL (New England Biolabs, Beverly, MA) and pRIT5 (Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein.

Purified fusion proteins can be utilized in VR-3 or VR-5 activity assays, (*e.g.*, direct assays or competitive assays described in detail below), or to generate antibodies specific for VR-3 or VR-5 proteins, for example. In a preferred embodiment, a VR-3 or VR-5 fusion protein expressed in a retroviral expression vector of the present invention can be utilized to infect bone marrow cells which are subsequently transplanted into

irradiated recipients. The pathology of the subject recipient is then examined after sufficient time has passed (*e.g.*, six (6) weeks).

Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amann *et al.*, (1988) *Gene* 69:301-315) and pET 11d (Studier *et al.*, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 60-89). Target gene expression from the pTrc vector relies on host RNA polymerase transcription from a hybrid trp-lac fusion promoter. Target gene expression from the pET 11d vector relies on transcription from a T7 gn10-lac fusion promoter mediated by a coexpressed viral RNA polymerase (T7 gn1). This viral polymerase is supplied by host strains BL21(DE3) or HMS174(DE3) from a resident prophage harboring a T7 gn1 gene under the transcriptional control of the lacUV 5 promoter.

One strategy to maximize recombinant protein expression in *E. coli* is to express the protein in a host bacteria with an impaired capacity to proteolytically cleave the recombinant protein (Gottesman, S., *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 119-128). Another strategy is to alter the nucleic acid sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in *E. coli* (Wada *et al.*, (1992) *Nucleic Acids Res.* 20:2111-2118). Such alteration of nucleic acid sequences of the invention can be carried out by standard DNA synthesis techniques.

In another embodiment, the VR-3 or VR-5 expression vector is a yeast expression vector. Examples of vectors for expression in yeast *S. cerevisiae* include pYepSec1 (Baldari, *et al.*, (1987) *Embo J.* 6:229-234), pMFa (Kurjan and Herskowitz, (1982) *Cell* 30:933-943), pJRY88 (Schultz *et al.*, (1987) *Gene* 54:113-123), pYES2 (Invitrogen Corporation, San Diego, CA), and picZ (Invitrogen Corp, San Diego, CA).

Alternatively, VR-3 or VR-5 proteins can be expressed in insect cells using baculovirus expression vectors. Baculovirus vectors available for expression of proteins in cultured insect cells (*e.g.*, Sf9 cells) include the pAc series (Smith *et al.* (1983) *Mol. Cell Biol.* 3:2156-2165) and the pVL series (Lucklow and Summers (1989) *Virology* 170:31-39).

In yet another embodiment, a nucleic acid of the invention is expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed, B. (1987) *Nature* 329:840) and pMT2PC (Kaufman *et al.* (1987) *EMBO J.* 6:187-195). When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40. For other suitable expression systems for both prokaryotic and eukaryotic cells see chapters 16 and 17 of Sambrook, J., Fritsch, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual*. 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

In another embodiment, the recombinant mammalian expression vector is capable of directing expression of the nucleic acid preferentially in a particular cell type (e.g., tissue-specific regulatory elements are used to express the nucleic acid). Tissue-specific regulatory elements are known in the art. Non-limiting examples of suitable tissue-specific promoters include the albumin promoter (liver-specific; Pinkert *et al.* (1987) *Genes Dev.* 1:268-277), lymphoid-specific promoters (Calame and Eaton (1988) *Adv. Immunol.* 43:235-275), in particular promoters of T cell receptors (Winoto and Baltimore (1989) *EMBO J.* 8:729-733) and immunoglobulins (Banerji *et al.* (1983) *Cell* 33:729-740; Queen and Baltimore (1983) *Cell* 33:741-748), neuron-specific promoters (e.g., the neurofilament promoter; Byrne and Ruddle (1989) *Proc. Natl. Acad. Sci. USA* 86:5473-5477), pancreas-specific promoters (Edlund *et al.* (1985) *Science* 230:912-916), and mammary gland-specific promoters (e.g., milk whey promoter; U.S. Patent No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are also encompassed, for example the murine hox promoters (Kessel and Gruss (1990) *Science* 249:374-379) and the  $\alpha$ -fetoprotein promoter (Campes and Tilghman (1989) *Genes Dev.* 3:537-546).

The expression characteristics of an endogenous VR-3 or VR-5 gene within a cell line or microorganism may be modified by inserting a heterologous DNA regulatory element into the genome of a stable cell line or cloned microorganism such that the inserted regulatory element is operatively linked with the endogenous VR-3 or VR-5 gene. For example, an endogenous VR-3 or VR-5 gene which is normally

“transcriptionally silent”, *i.e.*, a VR-3 or VR-5 gene which is normally not expressed, or is expressed only at very low levels in a cell line or microorganism, may be activated by inserting a regulatory element which is capable of promoting the expression of a normally expressed gene product in that cell line or microorganism. Alternatively, a transcriptionally silent, endogenous VR-3 or VR-5 gene may be activated by insertion of a promiscuous regulatory element that works across cell types.

A heterologous regulatory element may be inserted into a stable cell line or cloned microorganism, such that it is operatively linked with an endogenous VR-3 or VR-5 gene, using techniques, such as targeted homologous recombination, which are well known to those of skill in the art, and described, *e.g.*, in Chappel, U.S. Patent No. 5,272,071; PCT publication No. WO 91/06667, published May 16, 1991.

The invention further provides a recombinant expression vector comprising a DNA molecule of the invention cloned into the expression vector in an antisense orientation. That is, the DNA molecule is operatively linked to a regulatory sequence in a manner which allows for expression (by transcription of the DNA molecule) of an RNA molecule which is antisense to VR-3 or VR-5 mRNA. Regulatory sequences operatively linked to a nucleic acid cloned in the antisense orientation can be chosen which direct the continuous expression of the antisense RNA molecule in a variety of cell types, for instance viral promoters and/or enhancers, or regulatory sequences can be chosen which direct constitutive, tissue specific or cell type specific expression of antisense RNA. The antisense expression vector can be in the form of a recombinant plasmid, phagemid or attenuated virus in which antisense nucleic acids are produced under the control of a high efficiency regulatory region, the activity of which can be determined by the cell type into which the vector is introduced. For a discussion of the regulation of gene expression using antisense genes see Weintraub, H. *et al.*, Antisense RNA as a molecular tool for genetic analysis, *Reviews - Trends in Genetics*, Vol. 1(1) 1986.

Another aspect of the invention pertains to host cells into which a VR-3 or VR-5 nucleic acid molecule of the invention is introduced, *e.g.*, a VR-3 or VR-5 nucleic acid molecule within a recombinant expression vector or a VR-3 or VR-5 nucleic acid molecule containing sequences which allow it to homologously recombine into a specific site of the host cell's genome. The terms "host cell" and "recombinant host cell"

are used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

A host cell can be any prokaryotic or eukaryotic cell. For example, a VR-3 or VR-5 protein can be expressed in bacterial cells such as *E. coli*, insect cells, yeast or mammalian cells (such as Chinese hamster ovary cells (CHO) or COS cells). Other suitable host cells are known to those skilled in the art.

Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (*e.g.*, DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, or electroporation. Suitable methods for transforming or transfecting host cells can be found in Sambrook, *et al.* (*Molecular Cloning: A Laboratory Manual*, 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989), and other laboratory manuals.

For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of cells may integrate the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable marker (*e.g.*, resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Preferred selectable markers include those which confer resistance to drugs, such as G418, hygromycin and methotrexate. Nucleic acid encoding a selectable marker can be introduced into a host cell on the same vector as that encoding a VR-3 or VR-5 protein or can be introduced on a separate vector. Cells stably transfected with the introduced nucleic acid can be identified by drug selection (*e.g.*, cells that have incorporated the selectable marker gene will survive, while the other cells die).

A host cell of the invention, such as a prokaryotic or eukaryotic host cell in culture, can be used to produce (*i.e.*, express) a VR-3 or VR-5 protein. Accordingly, the invention further provides methods for producing a VR-3 or VR-5 protein using the host

cells of the invention. In one embodiment, the method comprises culturing the host cell of the invention (into which a recombinant expression vector encoding a VR-3 or VR-5 protein has been introduced) in a suitable medium such that a VR-3 or VR-5 protein is produced. In another embodiment, the method further comprises isolating a VR-3 or VR-5 protein from the medium or the host cell.

The host cells of the invention can also be used to produce non-human transgenic animals. For example, in one embodiment, a host cell of the invention is a fertilized oocyte or an embryonic stem cell into which VR-3- or VR-5-coding sequences have been introduced. Such host cells can then be used to create non-human transgenic animals in which exogenous VR-3 or VR-5 sequences have been introduced into their genome or homologous recombinant animals in which endogenous VR-3 or VR-5 sequences have been altered. Such animals are useful for studying the function and/or activity of a VR-3 or VR-5 and for identifying and/or evaluating modulators of VR-3 or VR-5 activity. As used herein, a "transgenic animal" is a non-human animal, preferably a mammal, more preferably a rodent such as a rat or mouse, in which one or more of the cells of the animal includes a transgene. Other examples of transgenic animals include non-human primates, sheep, dogs, cows, goats, chickens, amphibians, and the like. A transgene is exogenous DNA which is integrated into the genome of a cell from which a transgenic animal develops and which remains in the genome of the mature animal, thereby directing the expression of an encoded gene product in one or more cell types or tissues of the transgenic animal. As used herein, a "homologous recombinant animal" is a non-human animal, preferably a mammal, more preferably a mouse, in which an endogenous VR-3 or VR-5 gene has been altered by homologous recombination between the endogenous gene and an exogenous DNA molecule introduced into a cell of the animal, *e.g.*, an embryonic cell of the animal, prior to development of the animal.

A transgenic animal of the invention can be created by introducing a VR-3- or VR-5-encoding nucleic acid into the male pronuclei of a fertilized oocyte, *e.g.*, by microinjection, retroviral infection, and allowing the oocyte to develop in a pseudopregnant female foster animal. The VR-3 or VR-5 cDNA sequence of SEQ ID NO:1,3, 4, or 6 can be introduced as a transgene into the genome of a non-human animal. Alternatively, a nonhuman homologue of a human VR-3 or VR-5 gene, such as a mouse or rat VR-3 or VR-5 gene, can be used as a transgene. Alternatively, a VR-3 or

VR-5 gene homologue, such as another VR-3 or VR-5 family member, can be isolated based on hybridization to the VR-3 or VR-5 cDNA sequences of SEQ ID NO:1,3, 4, or 6, or the DNA insert of the plasmid deposited with ATCC as Accession Number PTA-2013 (described further in subsection I above) and used as a transgene. Intronic  
5 sequences and polyadenylation signals can also be included in the transgene to increase the efficiency of expression of the transgene. A tissue-specific regulatory sequence(s) can be operably linked to a VR-3 or VR-5 transgene to direct expression of a VR-3 or VR-5 protein to particular cells. Methods for generating transgenic animals via embryo manipulation and microinjection, particularly animals such as mice, have become  
10 conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866 and 4,870,009, both by Leder *et al.*, U.S. Patent No. 4,873,191 by Wagner *et al.* and in Hogan, B., *Manipulating the Mouse Embryo*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986). Similar methods are used for production of other transgenic animals. A transgenic founder animal can be identified based upon the  
15 presence of a VR-3 or VR-5 transgene in its genome and/or expression of VR-3 or VR-5 mRNA in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying a transgene encoding a VR-3 or VR-5 protein can further be bred to other transgenic animals carrying other transgenes.

20 To create a homologous recombinant animal, a vector is prepared which contains at least a portion of a VR-3 or VR-5 gene into which a deletion, addition or substitution has been introduced to thereby alter, *e.g.*, functionally disrupt, the VR-3 or VR-5 gene. The VR-3 or VR-5 gene can be a human gene (*e.g.*, the cDNA of SEQ ID NO:1,3, 4, or 6), but more preferably, is a non-human homologue of a human VR-3 or VR-5 gene  
25 (*e.g.*, a cDNA isolated by stringent hybridization with the nucleotide sequence of SEQ ID NO:1,3, 4, or 6). For example, a mouse VR-3 or VR-5 gene can be used to construct a homologous recombination nucleic acid molecule, *e.g.*, a vector, suitable for altering an endogenous VR-3 or VR-5 gene in the mouse genome.

In a preferred embodiment, the homologous recombination nucleic acid  
30 molecule is designed such that, upon homologous recombination, the endogenous VR-3 or VR-5 gene is functionally disrupted (*i.e.*, no longer encodes a functional protein; also referred to as a "knock out" vector). Alternatively, the homologous recombination

nucleic acid molecule can be designed such that, upon homologous recombination, the endogenous VR-3 or VR-5 gene is mutated or otherwise altered but still encodes functional protein (*e.g.*, the upstream regulatory region can be altered to thereby alter the expression of the endogenous VR-3 or VR-5 protein). In the homologous recombination

5 nucleic acid molecule, the altered portion of the VR-3 or VR-5 gene is flanked at its 5' and 3' ends by additional nucleic acid sequence of the VR-3 or VR-5 gene to allow for homologous recombination to occur between the exogenous VR-3 or VR-5 gene carried by the homologous recombination nucleic acid molecule and an endogenous VR-3 or VR-5 gene in a cell, *e.g.*, an embryonic stem cell. The additional flanking VR-3 or VR-

10 5 nucleic acid sequence is of sufficient length for successful homologous recombination with the endogenous gene. Typically, several kilobases of flanking DNA (both at the 5' and 3' ends) are included in the homologous recombination nucleic acid molecule (see, *e.g.*, Thomas, K.R. and Capecchi, M. R. (1987) *Cell* 51:503 for a description of homologous recombination vectors). The homologous recombination nucleic acid

15 molecule is introduced into a cell, *e.g.*, an embryonic stem cell line (*e.g.*, by electroporation) and cells in which the introduced VR-3 or VR-5 gene has homologously recombined with the endogenous VR-3 or VR-5 gene are selected (see *e.g.*, Li, E. *et al.* (1992) *Cell* 69:915). The selected cells can then be injected into a blastocyst of an animal (*e.g.*, a mouse) to form aggregation chimeras (see *e.g.*, Bradley, A. in *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, E.J. Robertson, ed. (IRL, Oxford, 1987) pp. 113-152). A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term. Progeny harboring the homologously recombined DNA in their germ cells can be used to breed animals in which all cells of the animal contain the homologously

20 recombined DNA by germline transmission of the transgene. Methods for constructing homologous recombination nucleic acid molecules, *e.g.*, vectors, or homologous recombinant animals are described further in Bradley, A. (1991) *Current Opinion in Biotechnology* 2:823-829 and in PCT International Publication Nos.: WO 90/11354 by Le Mouellec *et al.*; WO 91/01140 by Smithies *et al.*; WO 92/0968 by Zijlstra *et al.*; and

30 WO 93/04169 by Berns *et al.*



In another embodiment, transgenic non-human animals can be produced which contain selected systems which allow for regulated expression of the transgene. One example of such a system is the *cre/loxP* recombinase system of bacteriophage P1. For a description of the *cre/loxP* recombinase system, see, *e.g.*, Lakso *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:6232-6236. Another example of a recombinase system is the FLP recombinase system of *Saccharomyces cerevisiae* (O'Gorman *et al.* (1991) *Science* 251:1351-1355. If a *cre/loxP* recombinase system is used to regulate expression of the transgene, animals containing transgenes encoding both the *Cre* recombinase and a selected protein are required. Such animals can be provided through the construction of "double" transgenic animals, *e.g.*, by mating two transgenic animals, one containing a transgene encoding a selected protein and the other containing a transgene encoding a recombinase.

Clones of the non-human transgenic animals described herein can also be produced according to the methods described in Wilmut, I. *et al.* (1997) *Nature* 385:810-813 and PCT International Publication Nos. WO 97/07668 and WO 97/07669. In brief, a cell, *e.g.*, a somatic cell, from the transgenic animal can be isolated and induced to exit the growth cycle and enter G<sub>0</sub> phase. The quiescent cell can then be fused, *e.g.*, through the use of electrical pulses, to an enucleated oocyte from an animal of the same species from which the quiescent cell is isolated. The reconstructed oocyte is then cultured such that it develops to morula or blastocyte and then transferred to pseudopregnant female foster animal. The offspring borne of this female foster animal will be a clone of the animal from which the cell, *e.g.*, the somatic cell, is isolated.

#### IV. Pharmaceutical Compositions

The VR-3 or VR-5 nucleic acid molecules, fragments of VR-3 or VR-5 proteins, and anti-VR-3 or anti-VR-5 antibodies (also referred to herein as "active compounds") of the invention can be incorporated into pharmaceutical compositions suitable for administration. Such compositions typically comprise the nucleic acid molecule, protein, or antibody and a pharmaceutically acceptable carrier. As used herein the language "pharmaceutically acceptable carrier" is intended to include any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration.

The use of such media and agents for pharmaceutically active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active compound, use thereof in the compositions is contemplated. Supplementary active compounds can also be incorporated into the compositions.

5           A pharmaceutical composition of the invention is formulated to be compatible with its intended route of administration. Examples of routes of administration include parenteral, *e.g.*, intravenous, intradermal, subcutaneous, oral (*e.g.*, inhalation), transdermal (topical), transmucosal, and rectal administration. Solutions or suspensions used for parenteral, intradermal, or subcutaneous application can include the following  
10 components: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid; buffers such as acetates, citrates or phosphates and agents for the adjustment of  
15 tonicity such as sodium chloride or dextrose. pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic.

          Pharmaceutical compositions suitable for injectable use include sterile aqueous  
20 solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor EL™ (BASF, Parsippany, NJ) or phosphate buffered saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy  
25 syringability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), and suitable mixtures thereof. The proper fluidity  
30 can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various

- 61 -

antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as manitol, sorbitol, sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

Sterile injectable solutions can be prepared by incorporating the active compound (*e.g.*, a fragment of a VR-3 or VR-5 protein or an anti-VR-3 or anti-VR-5 antibody) in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle which contains a basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum drying and freeze-drying which yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

Oral compositions generally include an inert diluent or an edible carrier. They can be enclosed in gelatin capsules or compressed into tablets. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients and used in the form of tablets, troches, or capsules. Oral compositions can also be prepared using a fluid carrier for use as a mouthwash, wherein the compound in the fluid carrier is applied orally and swished and expectorated or swallowed. Pharmaceutically compatible binding agents, and/or adjuvant materials can be included as part of the composition. The tablets, pills, capsules, troches and the like can contain any of the following ingredients, or compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin; an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel, or corn starch; a lubricant such as magnesium stearate or Sterotes; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate, or orange flavoring.

For administration by inhalation, the compounds are delivered in the form of an aerosol spray from pressured container or dispenser which contains a suitable propellant, *e.g.*, a gas such as carbon dioxide, or a nebulizer.

Systemic administration can also be by transmucosal or transdermal means. For  
5 transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active  
10 compounds are formulated into ointments, salves, gels, or creams as generally known in the art.

The compounds can also be prepared in the form of suppositories (*e.g.*, with conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

15 In one embodiment, the active compounds are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid.  
20 Methods for preparation of such formulations will be apparent to those skilled in the art. The materials can also be obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes targeted to infected cells with monoclonal antibodies to viral antigens) can also be used as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled  
25 in the art, for example, as described in U.S. Patent No. 4,522,811.

It is especially advantageous to formulate oral or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound  
30 calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier. The specification for the dosage unit forms of the invention are dictated by and directly dependent on the unique characteristics of the active compound

and the particular therapeutic effect to be achieved, and the limitations inherent in the art of compounding such an active compound for the treatment of individuals.

Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g.*, for  
5 determining the LD50 (the dose lethal to 50% of the population) and the ED50 (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD50/ED50. Compounds which exhibit large therapeutic indices are preferred. While compounds that exhibit toxic side effects may be used, care should be taken to design a  
10 delivery system that targets such compounds to the site of affected tissue in order to minimize potential damage to uninfected cells and, thereby, reduce side effects.

The data obtained from the cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED50 with little  
15 or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range that includes the IC50 (*i.e.*, the concentration of  
20 the test compound which achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by high performance liquid chromatography.

As defined herein, a therapeutically effective amount of protein or polypeptide  
25 (*i.e.*, an effective dosage) ranges from about 0.001 to 30 mg/kg body weight, preferably about 0.01 to 25 mg/kg body weight, more preferably about 0.1 to 20 mg/kg body weight, and even more preferably about 1 to 10 mg/kg, 2 to 9 mg/kg, 3 to 8 mg/kg, 4 to 7 mg/kg, or 5 to 6 mg/kg body weight. The skilled artisan will appreciate that certain factors may influence the dosage required to effectively treat a subject, including but not  
30 limited to the severity of the disease or disorder, previous treatments, the general health and/or age of the subject, and other diseases present. Moreover, treatment of a subject

with a therapeutically effective amount of a protein, polypeptide, or antibody can include a single treatment or, preferably, can include a series of treatments.

In a preferred example, a subject is treated with antibody, protein, or polypeptide in the range of between about 0.1 to 20 mg/kg body weight, one time per week for  
5 between about 1 to 10 weeks, preferably between 2 to 8 weeks, more preferably between about 3 to 7 weeks, and even more preferably for about 4, 5, or 6 weeks. It will also be appreciated that the effective dosage of antibody, protein, or polypeptide used for treatment may increase or decrease over the course of a particular treatment. Changes in dosage may result and become apparent from the results of diagnostic assays as  
10 described herein.

The present invention encompasses agents which modulate expression or activity. An agent may, for example, be a small molecule. For example, such small molecules include, but are not limited to, peptides, peptidomimetics, amino acids, amino acid analogs, polynucleotides, polynucleotide analogs, nucleotides, nucleotide analogs,  
15 organic or inorganic compounds (*i.e.*, including heteroorganic and organometallic compounds) having a molecular weight less than about 10,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 5,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 1,000 grams per mole, organic or inorganic compounds having a molecular weight less than  
20 about 500 grams per mole, and salts, esters, and other pharmaceutically acceptable forms of such compounds. It is understood that appropriate doses of small molecule agents depends upon a number of factors within the ken of the ordinarily skilled physician, veterinarian, or researcher. The dose(s) of the small molecule will vary, for example, depending upon the identity, size, and condition of the subject or sample being  
25 treated, further depending upon the route by which the composition is to be administered, if applicable, and the effect which the practitioner desires the small molecule to have upon the nucleic acid or polypeptide of the invention.

Exemplary doses include milligram or microgram amounts of the small molecule per kilogram of subject or sample weight (*e.g.*, about 1 microgram per kilogram to about  
30 500 milligrams per kilogram, about 100 micrograms per kilogram to about 5 milligrams per kilogram, or about 1 microgram per kilogram to about 50 micrograms per kilogram. It is

furthermore understood that appropriate doses of a small molecule depend upon the potency of the small molecule with respect to the expression or activity to be modulated. Such appropriate doses may be determined using the assays described herein. When one or more of these small molecules is to be administered to an animal (*e.g.*, a human) in order to modulate expression or activity of a polypeptide or nucleic acid of the invention, a physician, veterinarian, or researcher may, for example, prescribe a relatively low dose at first, subsequently increasing the dose until an appropriate response is obtained. In addition, it is understood that the specific dose level for any particular animal subject will depend upon a variety of factors including the activity of the specific compound employed, the age, body weight, general health, gender, and diet of the subject, the time of administration, the route of administration, the rate of excretion, any drug combination, and the degree of expression or activity to be modulated.

Further, an antibody (or fragment thereof) may be conjugated to a therapeutic moiety such as a cytotoxin, a therapeutic agent or a radioactive metal ion. A cytotoxin or cytotoxic agent includes any agent that is detrimental to cells. Examples include taxol, cytochalasin B, gramicidin D, ethidium bromide, emetine, mitomycin, etoposide, teniposide, vincristine, vinblastine, colchicin, doxorubicin, daunorubicin, dihydroxy anthracin dione, mitoxantrone, mithramycin, actinomycin D, 1-dehydrotestosterone, glucocorticoids, procaine, tetracaine, lidocaine, propranolol, and puromycin and analogs or homologs thereof. Therapeutic agents include, but are not limited to, antimetabolites (*e.g.*, methotrexate, 6-mercaptopurine, 6-thioguanine, cytarabine, 5-fluorouracil decarbazine), alkylating agents (*e.g.*, mechlorethamine, thioepa chlorambucil, melphalan, carmustine (BSNU) and lomustine (CCNU), cyclophosphamide, busulfan, dibromomannitol, streptozotocin, mitomycin C, and cis-dichlorodiamine platinum (II) (DDP) cisplatin), anthracyclines (*e.g.*, daunorubicin (formerly daunomycin) and doxorubicin), antibiotics (*e.g.*, dactinomycin (formerly actinomycin), bleomycin, mithramycin, and anthramycin (AMC)), and anti-mitotic agents (*e.g.*, vincristine and vinblastine).

The conjugates of the invention can be used for modifying a given biological response, the drug moiety is not to be construed as limited to classical chemical therapeutic agents. For example, the drug moiety may be a protein or polypeptide

possessing a desired biological activity. Such proteins may include, for example, a toxin such as abrin, ricin A, pseudomonas exotoxin, or diphtheria toxin; a protein such as tumor necrosis factor, alpha.-interferon, beta.-interferon, nerve growth factor, platelet derived growth factor, tissue plasminogen activator; or, biological response modifiers  
5 such as, for example, lymphokines, interleukin-1 ("IL-1"), interleukin-2 ("IL-2"), interleukin-6 ("IL-6"), granulocyte macrophase colony stimulating factor ("GM-CSF"), granulocyte colony stimulating factor ("G-CSF"), or other growth factors.

Techniques for conjugating such therapeutic moiety to antibodies are well known, see, *e.g.*, Arnon et al., "Monoclonal Antibodies For Immunotargeting Of Drugs  
10 In Cancer Therapy", in *Monoclonal Antibodies And Cancer Therapy*, Reisfeld et al. (eds.), pp. 243-56 (Alan R. Liss, Inc. 1985); Hellstrom et al., "Antibodies For Drug Delivery", in *Controlled Drug Delivery* (2nd Ed.), Robinson et al. (eds.), pp. 623-53 (Marcel Dekker, Inc. 1987); Thorpe, "Antibody Carriers Of Cytotoxic Agents In Cancer Therapy: A Review", in *Monoclonal Antibodies '84: Biological And Clinical  
15 Applications*, Pinchera et al. (eds.), pp. 475-506 (1985); "Analysis, Results, And Future Prospective Of The Therapeutic Use Of Radiolabeled Antibody In Cancer Therapy", in *Monoclonal Antibodies For Cancer Detection And Therapy*, Baldwin et al. (eds.), pp. 303-16 (Academic Press 1985), and Thorpe et al., "The Preparation And Cytotoxic Properties Of Antibody-Toxin Conjugates", *Immunol. Rev.*, 62:119-58 (1982).  
20 Alternatively, an antibody can be conjugated to a second antibody to form an antibody heteroconjugate as described by Segal in U.S. Patent No. 4,676,980.

The nucleic acid molecules of the invention can be inserted into vectors and used as gene therapy vectors. Gene therapy vectors can be delivered to a subject by, for example, intravenous injection, local administration (see U.S. Patent 5,328,470) or by  
25 stereotactic injection (see *e.g.*, Chen *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:3054-3057). The pharmaceutical preparation of the gene therapy vector can include the gene therapy vector in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery vector can be produced intact from recombinant cells, *e.g.*, retroviral vectors, the  
30 pharmaceutical preparation can include one or more cells which produce the gene delivery system.



The pharmaceutical compositions can be included in a container, pack, or dispenser together with instructions for administration.

#### V. Uses and Methods of the Invention

- 5           The nucleic acid molecules, proteins, protein homologues, and antibodies described herein can be used in one or more of the following methods: a) screening assays; b) predictive medicine (*e.g.*, diagnostic assays, prognostic assays, monitoring clinical trials, and pharmacogenetics); and c) methods of treatment (*e.g.*, therapeutic and prophylactic). As described herein, a VR-3 or VR-5 protein of the invention has one or
- 10 more of the following activities: (1) it interacts with a non-VR-3 or non-VR-5 protein molecule, *e.g.*, a VR-3 or VR-5 ligand such as a vanilloid compound, *e.g.*, capsaicin; (2) it activates a VR-3- or VR-5-dependent signal transduction pathway; (3) it modulates intracellular calcium concentration; (4) it modulates pain signaling mechanisms and/or calcium homeostasis mechanisms; and (5) it modulates cellular growth and/or
- 15 proliferation, and, thus, can be used to, for example, (1) modulate the interaction with a non-VR-3 or non-VR-5 protein molecule; (2) activate a VR-3- or VR-5-dependent signal transduction pathway; (3) modulate intracellular calcium concentrations; (4) modulate pain signaling mechanisms; (5) participate in nociception; (6) modulate cellular growth and proliferation disorders, *e.g.*, cancer; and (7) modulate tumorigenesis.
- 20           The isolated nucleic acid molecules of the invention can be used, for example, to express VR-3 or VR-5 protein (*e.g.*, via a recombinant expression vector in a host cell in gene therapy applications), to detect VR-3 or VR-5 mRNA (*e.g.*, in a biological sample) or a genetic alteration in a VR-3 or VR-5 gene, and to modulate VR-3 or VR-5 activity, as described further below. The VR-3 or VR-5 proteins can be used to treat disorders
- 25 characterized by insufficient or excessive production of a VR-3 or VR-5 ligand or substrate or production of VR-3 or VR-5 inhibitors. In addition, the VR-3 or VR-5 proteins can be used to screen for naturally occurring VR-3 or VR-5 ligands or substrates to screen for drugs or compounds which modulate VR-3 or VR-5 activity, as well as to treat disorders characterized by insufficient or excessive production of VR-3
- 30 or VR-5 protein or production of VR-3 or VR-5 protein forms which have decreased, aberrant or unwanted activity compared to VR-3 or VR-5 wild type protein (*e.g.*, calcium homeostasis related disorders, pain disorders, and/or cellular growth and/or

proliferation disorders, *e.g.*, cancer). Moreover, the anti-VR-3 or anti-VR-5 antibodies of the invention can be used to detect and isolate VR-3 or VR-5 proteins, regulate the bioavailability of VR-3 or VR-5 proteins, and modulate VR-3 or VR-5 activity.

5                   A.     Screening Assays:

The invention provides a method (also referred to herein as a "screening assay") for identifying modulators, *i.e.*, candidate or test compounds or agents (*e.g.*, peptides, peptidomimetics, small molecules or other drugs) which bind to VR-3 or VR-5 proteins, have a stimulatory or inhibitory effect on, for example, VR-3 or VR-5 expression or  
10 VR-3 or VR-5 activity, or have a stimulatory or inhibitory effect on, for example, the expression or activity of a VR-3 or VR-5 ligand or substrate.

In one embodiment, the invention provides assays for screening candidate or test compounds which are substrates or ligands of a VR-3 or VR-5 protein or polypeptide or biologically active portion thereof. In another embodiment, the invention provides  
15 assays for screening candidate or test compounds which bind to or modulate the activity of a VR-3 or VR-5 protein or polypeptide or biologically active portion thereof. The test compounds of the present invention can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; spatially addressable parallel solid phase or solution phase libraries; synthetic  
20 library methods requiring deconvolution; the 'one-bead one-compound' library method; and synthetic library methods using affinity chromatography selection. The biological library approach is limited to peptide libraries, while the other four approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam, K.S. (1997) *Anticancer Drug Des.* 12:145).

25           Examples of methods for the synthesis of molecular libraries can be found in the art, for example in: DeWitt *et al.* (1993) *Proc. Natl. Acad. Sci. U.S.A.* 90:6909; Erb *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:11422; Zuckermann *et al.* (1994) *J. Med. Chem.* 37:2678; Cho *et al.* (1993) *Science* 261:1303; Carrell *et al.* (1994) *Angew. Chem. Int. Ed. Engl.* 33:2059; Carell *et al.* (1994) *Angew. Chem. Int. Ed. Engl.*  
30 33:2061; and in Gallop *et al.* (1994) *J. Med. Chem.* 37:1233.

Libraries of compounds may be presented in solution (*e.g.*, Houghten (1992) *Biotechniques* 13:412-421), or on beads (Lam (1991) *Nature* 354:82-84), chips (Fodor (1993) *Nature* 364:555-556), bacteria (Ladner USP 5,223,409), spores (Ladner USP '409), plasmids (Cull *et al.* (1992) *Proc Natl Acad Sci USA* 89:1865-1869) or on phage  
5 (Scott and Smith (1990) *Science* 249:386-390); (Devlin (1990) *Science* 249:404-406); (Cwirla *et al.* (1990) *Proc. Natl. Acad. Sci.* 87:6378-6382); (Felici (1991) *J. Mol. Biol.* 222:301-310); (Ladner *supra.*).

In one embodiment, an assay is a cell-based assay in which a cell which expresses a VR-3 or VR-5 protein or biologically active portion thereof is contacted  
10 with a test compound and the ability of the test compound to modulate VR-3 or VR-5 activity is determined. Determining the ability of the test compound to modulate VR-3 or VR-5 activity can be accomplished by monitoring, for example, intracellular calcium concentration or membrane depolarization by, *e.g.*, patch-clamp recordings in whole-cell, inside-out, and outside-out configurations (as described in, for example, Tominaga  
15 M. *et al.* (1998) *Neuron* 21:531-543), or the activity of a VR-3- or VR-5-regulated transcription factor. The cell, for example, can be of mammalian origin, *e.g.*, a neuronal cell.

High throughput screens may also be used to detect the ability of the test compound to modulate VR-3 or VR-5 activity. High throughput screens may include  
20 fluorescence based assays using the Fluorometric Imaging Plate Reader (FLIPR) with calcium sensitive dyes, and reporter gene assays using calcium sensitive photoproteins that emit light on the influx of calcium and can be detected using an Imaging system. Determining the ability of the test compound to modulate VR-3 or VR-5 activity can also be accomplished by monitoring, for example, pain signaling mechanisms.

25 The ability of the test compound to modulate VR-3 or VR-5 binding to a ligand or substrate or to bind to VR-3 or VR-5 can also be determined. Determining the ability of the test compound to modulate VR-3 or VR-5 binding to a ligand or substrate can be accomplished, for example, by coupling the VR-3 or VR-5 ligand or substrate with a radioisotope or enzymatic label such that binding of the VR-3 or VR-5 ligand or  
30 substrate to VR-3 or VR-5 can be determined by detecting the labeled VR-3 or VR-5 ligand or substrate in a complex. Determining the ability of the test compound to bind VR-3 or VR-5 can be accomplished, for example, by coupling the compound with a

radioisotope or enzymatic label such that binding of the compound to VR-3 or VR-5 can be determined by detecting the labeled VR-3 or VR-5 compound in a complex. For example, compounds (*e.g.*, VR-3 or VR-5 ligands or substrates, *e.g.*, capsaicin) can be labeled with  $^{125}\text{I}$ ,  $^{35}\text{S}$ ,  $^{14}\text{C}$ , or  $^3\text{H}$ , either directly or indirectly, and the radioisotope  
5 detected by direct counting of radioemmission or by scintillation counting.

Alternatively, compounds can be enzymatically labeled with, for example, horseradish peroxidase, alkaline phosphatase, or luciferase, and the enzymatic label detected by determination of conversion of an appropriate substrate to product.

It is also within the scope of this invention to determine the ability of a  
10 compound (*e.g.*, a VR-3 or VR-5 ligand or substrate) to interact with VR-3 or VR-5 without the labeling of any of the interactants. For example, a microphysiometer can be used to detect the interaction of a compound with VR-3 or VR-5 without the labeling of either the compound or the VR-3 or VR-5. McConnell, H. M. *et al.* (1992) *Science* 257:1906-1912. As used herein, a "microphysiometer" (*e.g.*, Cytosensor) is an  
15 analytical instrument that measures the rate at which a cell acidifies its environment using a light-addressable potentiometric sensor (LAPS). Changes in this acidification rate can be used as an indicator of the interaction between a compound and VR-3 or VR-5.

In another embodiment, an assay is a cell-based assay comprising contacting a  
20 cell expressing a VR-3 or VR-5 target molecule (*e.g.*, a VR-3 or VR-5 ligand or substrate) with a test compound and determining the ability of the test compound to modulate (*e.g.*, stimulate or inhibit) the activity of the VR-3 or VR-5 target molecule. Determining the ability of the test compound to modulate the activity of a VR-3 or VR-5 target molecule can be accomplished, for example, by determining the ability of the VR-  
25 3 or VR-5 protein to bind to or interact with the VR-3 or VR-5 target molecule.

Determining the ability of the VR-3 or VR-5 protein or a biologically active fragment thereof, to bind to or interact with a VR-3 or VR-5 target molecule can be accomplished by one of the methods described above for determining direct binding. In a preferred embodiment, determining the ability of the VR-3 or VR-5 protein to bind to  
30 or interact with a VR-3 or VR-5 target molecule can be accomplished by determining the activity of the target molecule. For example, the activity of the target molecule can be determined by detecting induction of a cellular second messenger of the target (*i.e.*,

intracellular  $\text{Ca}^{2+}$ , diacylglycerol,  $\text{IP}_3$ , and the like), detecting catalytic/enzymatic activity of the target an appropriate substrate, detecting the induction of a reporter gene (comprising a target-responsive regulatory element operatively linked to a nucleic acid encoding a detectable marker, *e.g.*, luciferase), or detecting a target-regulated cellular response such as changes in membrane permeability to ions, *e.g.*, changes in membrane potential, or changes in intracellular calcium levels (as measured, for example, by flow cytometry).

In yet another embodiment, an assay of the present invention is a cell-free assay in which a VR-3 or VR-5 protein or biologically active portion thereof is contacted with a test compound and the ability of the test compound to bind to the VR-3 or VR-5 protein or biologically active portion thereof is determined. Preferred biologically active portions of the VR-3 or VR-5 proteins to be used in assays of the present invention include fragments which participate in interactions with non-VR-3 or anti-VR-5 molecules, *e.g.*, fragments with high surface probability scores (see, for example, Figures 6 and 7). Binding of the test compound to the VR-3 or VR-5 protein can be determined either directly or indirectly as described above. In a preferred embodiment, the assay includes contacting the VR-3 or VR-5 protein or biologically active portion thereof with a known compound which binds VR-3 or VR-5 to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a VR-3 or VR-5 protein, wherein determining the ability of the test compound to interact with a VR-3 or VR-5 protein comprises determining the ability of the test compound to preferentially bind to VR-3 or VR-5 or biologically active portion thereof as compared to the known compound.

In another embodiment, the assay is a cell-free assay in which a VR-3 or VR-5 protein or biologically active portion thereof is contacted with a test compound and the ability of the test compound to modulate (*e.g.*, stimulate or inhibit) the activity of the VR-3 or VR-5 protein or biologically active portion thereof is determined. Determining the ability of the test compound to modulate the activity of a VR-3 or VR-5 protein can be accomplished, for example, by determining the ability of the VR-3 or VR-5 protein to bind to a VR-3 or VR-5 target molecule by one of the methods described above for determining direct binding. Determining the ability of the VR-3 or VR-5 protein to bind to a VR-3 or VR-5 target molecule can also be accomplished using a technology such as

real-time Biomolecular Interaction Analysis (BIA). Sjolander, S. and Urbaniczky, C. (1991) *Anal. Chem.* 63:2338-2345 and Szabo *et al.* (1995) *Curr. Opin. Struct. Biol.* 5:699-705. As used herein, "BIA" is a technology for studying biospecific interactions in real time, without labeling any of the interactants (*e.g.*, BIAcore). Changes in the optical phenomenon of surface plasmon resonance (SPR) can be used as an indication of real-time reactions between biological molecules.

In an alternative embodiment, determining the ability of the test compound to modulate the activity of a VR-3 or VR-5 protein can be accomplished by determining the ability of the VR-3 or VR-5 protein to further modulate the activity of a downstream effector of a VR-3 or VR-5 target molecule. For example, the activity of the effector molecule on an appropriate target can be determined or the binding of the effector to an appropriate target can be determined as previously described.

In yet another embodiment, the cell-free assay involves contacting a VR-3 or VR-5 protein or biologically active portion thereof with a known compound which binds the VR-3 or VR-5 protein to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with the VR-3 or VR-5 protein, wherein determining the ability of the test compound to interact with the VR-3 or VR-5 protein comprises determining the ability of the VR-3 or VR-5 protein to preferentially bind to or modulate the activity of a VR-3 or VR-5 target molecule.

In more than one embodiment of the above assay methods of the present invention, it may be desirable to immobilize either VR-3 or VR-5 or its target molecule to facilitate separation of complexed from uncomplexed forms of one or both of the proteins, as well as to accommodate automation of the assay. Binding of a test compound to a VR-3 or VR-5 protein, or interaction of a VR-3 or VR-5 protein with a target molecule in the presence and absence of a candidate compound, can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtitre plates, test tubes, and micro-centrifuge tubes. In one embodiment, a fusion protein can be provided which adds a domain that allows one or both of the proteins to be bound to a matrix. For example, glutathione-S-transferase/VR-3 or VR-5 fusion proteins or glutathione-S-transferase/target fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or

glutathione derivatized microtitre plates, which are then combined with the test compound or the test compound and either the non-adsorbed target protein or VR-3 or VR-5 protein, and the mixture incubated under conditions conducive to complex formation (*e.g.*, at physiological conditions for salt and pH). Following incubation, the  
5 beads or microtitre plate wells are washed to remove any unbound components, the matrix immobilized in the case of beads, complex determined either directly or indirectly, for example, as described above. Alternatively, the complexes can be dissociated from the matrix, and the level of VR-3 or VR-5 binding or activity determined using standard techniques.

10 Other techniques for immobilizing proteins on matrices can also be used in the screening assays of the invention. For example, either a VR-3 or VR-5 protein or a VR-3 or VR-5 target molecule can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated VR-3 or VR-5 protein or target molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques known in the art (*e.g.*,  
15 biotinylation kit, Pierce Chemicals, Rockford, IL), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies reactive with VR-3 or VR-5 protein or target molecules but which do not interfere with binding of the VR-3 or VR-5 protein to its target molecule can be derivatized to the wells of the plate, and unbound target or VR-3 or VR-5 protein trapped in the wells by antibody  
20 conjugation. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the VR-3 or VR-5 protein or target molecule, as well as enzyme-linked assays which rely on detecting an enzymatic activity associated with the VR-3 or VR-5 protein or target molecule.

25 In another embodiment, modulators of VR-3 or VR-5 expression are identified in a method wherein a cell is contacted with a candidate compound and the expression of VR-3 or VR-5 mRNA or protein in the cell is determined. The level of expression of VR-3 or VR-5 mRNA or protein in the presence of the candidate compound is compared to the level of expression of VR-3 or VR-5 mRNA or protein in the absence of the  
30 candidate compound. The candidate compound can then be identified as a modulator of VR-3 or VR-5 expression based on this comparison. For example, when expression of VR-3 or VR-5 mRNA or protein is greater (statistically significantly greater) in the

presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of VR-3 or VR-5 mRNA or protein expression. Alternatively, when expression of VR-3 or VR-5 mRNA or protein is less (statistically significantly less) in the presence of the candidate compound than in its absence, the candidate  
5 compound is identified as an inhibitor of VR-3 or VR-5 mRNA or protein expression. The level of VR-3 or VR-5 mRNA or protein expression in the cells can be determined by methods described herein for detecting VR-3 or VR-5 mRNA or protein.

In yet another aspect of the invention, the VR-3 or VR-5 proteins can be used as "bait proteins" in a two-hybrid assay or three-hybrid assay (see, *e.g.*, U.S. Patent No. 5,283,317; Zervos *et al.* (1993) *Cell* 72:223-232; Madura *et al.* (1993) *J. Biol. Chem.* 268:12046-12054; Bartel *et al.* (1993) *Biotechniques* 14:920-924; Iwabuchi *et al.* (1993) *Oncogene* 8:1693-1696; and Brent WO94/10300), to identify other proteins, which bind to or interact with VR-3 or VR-5 ("VR-3-binding proteins" or "VR-3-bp" or "VR-5-binding proteins" or "VR-5-bp") and are involved in VR-3 or VR-5 activity.  
15 Such VR-3- or VR-5-binding proteins are also likely to be involved in the propagation of signals by the VR-3 or VR-5 proteins or VR-3 or VR-5 targets as, for example, downstream elements of a VR-3- or VR5-mediated signaling pathway. Alternatively, such VR-3- or VR5-binding proteins are likely to be VR-3 or VR-5 inhibitors.

The two-hybrid system is based on the modular nature of most transcription  
20 factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. In one construct, the gene that codes for a VR-3 or VR-5 protein is fused to a gene encoding the DNA binding domain of a known transcription factor (*e.g.*, GAL-4). In the other construct, a DNA sequence, from a library of DNA sequences, that encodes an unidentified protein ("prey" or "sample") is  
25 fused to a gene that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact, *in vivo*, forming a VR-3- or VR5-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (*e.g.*, LacZ) which is operably linked to a transcriptional regulatory site responsive to  
30 the transcription factor. Expression of the reporter gene can be detected and cell colonies containing the functional transcription factor can be isolated and used to obtain



- 75 -

the cloned gene which encodes the protein which interacts with the VR-3 or VR-5 protein.

In another aspect, the invention pertains to a combination of two or more of the assays described herein. For example, a modulating agent can be identified using a cell-based or a cell free assay, and the ability of the agent to modulate the activity of a VR-3 or VR-5 protein can be confirmed *in vivo*, *e.g.*, in an animal such as an animal model for pain or an animal model for a cellular growth or proliferation disorder, *e.g.*, cancer.

Models for studying pain *in vivo* include, but are not limited to, rat models of neuropathic pain caused by methods such as intraperitoneal administration of Taxol (Authier *et al.* (2000) *Brain Res.* 887:239-249), chronic constriction injury (CCI), partial sciatic nerve transection (Linenlaub and Sommer (2000) *Pain* 89:97-106), transection of the tibial and sural nerves (Lee *et al.* (2000) *Neurosci. Lett.* 291:29-32), the spared nerve injury model (Decosterd and Woolf (2000) *Pain* 87:149-158), cuffing the sciatic nerve (Pitcher and Henry (2000) *Eur. J. Neurosci.* 12:2006-2020), unilateral tight ligation (Esser and Sawynok (2000) *Eur. J. Pharmacol.* 399:131-139), L5 spinal nerve ligation (Honroe *et al.* (2000) *Neurosci.* 98:585-598), and photochemically induced ischemic nerve injury (Hao *et al.* (2000) *Exp. Neurol.* 163:231-238); rat models of nociceptive pain caused by methods such as the Chung Method, the Bennett Method, and intraperitoneal administration of inflammatory agents such as carageenan, formalin, and complete Freund's adjuvant (CFA) (Abdi *et al.* (2000) *Anesth. Analg.* 91:955-959); rat models of post-incisional pain caused by incising the skin and fascia of a hind paw (Olivera and Prado (2000) *Braz. J. Med. Biol. Res.* 33:957-960); rat models of cancer pain caused by methods such as injecting osteolytic sarcoma cells into the femur (Honroe *et al.* (2000) *Neurosci.* 98:585-598); and rat models of visceral pain caused by methods such as intraperitoneal administration of cyclophosphamide. Other screens may involve the study of modulators in human volunteers subject to topically applied capsaicin.

Various methods of determining an animal's response to pain are known in the art. Examples of such methods include, but are not limited to, brief intense exposure to a focused heat source, administration of a noxious chemical subcutaneously, the tail flick test, the hot plate test, the formalin test, Von Frey threshold, and testing for stress-

induced analgesia (*et al.*, by restraint, foot shock, and/or cold water swim) (Crawley (2000) *What's Wrong With My Mouse?* Wiley-Liss pp. 72-75).

Examples of animal models of cancer include transplantable models (*e.g.*, xenografts of colon tumors such as Co-3, AC3603 or WiDr or into  
5 immunocompromised mice such as SCID or nude mice); transgenic models (*e.g.*, B66-Min/+ mouse); chemical induction models, *e.g.*, carcinogen (*e.g.*, azoxymethane, 2-dimethylhydrazine, or N-nitrosodimethylamine) treated rats or mice; models of liver metastasis from colon cancer such as that described by Rashidi *et al.* (2000) *Anticancer Res* 20(2A):715; rodent models of breast cancer such as that described by Blackshear  
10 P.E. (2001) *Toxicol Pathol* 29(1):105-16; and cancer cell implantation or inoculation models as described in, for example, Fingert, *et al.* (1987) *Cancer Res* 46(14):3824-9 and Teraoka, *et al.* (1995) *Jpn J Cancer Res* 86(5):419-23.

This invention further pertains to novel agents identified by the above-described screening assays. Accordingly, it is within the scope of this invention to further use an  
15 agent identified as described herein in an appropriate animal model. For example, an agent identified as described herein (*e.g.*, a VR-3 or VR-5 modulating agent, an antisense VR-3 or VR-5 nucleic acid molecule, a VR-3 or VR-5-specific antibody, or a VR-3 or VR-5-binding partner) can be used in an animal model to determine the efficacy, toxicity, or side effects of treatment with such an agent. Alternatively, an  
20 agent identified as described herein can be used in an animal model to determine the mechanism of action of such an agent. Furthermore, this invention pertains to uses of novel agents identified by the above-described screening assays for treatments as described herein. In one embodiment, the invention features a method of treating a subject having a cellular growth or proliferation disorder that involves administering to  
25 the subject an VR-3 or VR-5 modulator such that treatment occurs. In another embodiment, the invention features a method of treating a subject having cancer, *e.g.*, colon, breast, prostate, lung, or ovarian cancer, that involves treating a subject with an VR-3 or VR-5 modulator, such that treatment occurs. Preferred VR-3 or VR-5 modulators include, but are not limited to, VR-3 or VR-5 proteins or biologically active  
30 fragments, VR-3 or VR-5 nucleic acid molecules, VR-3 or VR-5 antibodies, ribozymes, and VR-3 or VR-5 antisense oligonucleotides designed based on the VR-3 or VR-5 nucleotide sequences disclosed herein, as well as peptides, organic and non-organic

small molecules identified as being capable of modulating VR-3 or VR-5 expression and/or activity, for example, according to at least one of the screening assays described herein.

Any of the compounds, including but not limited to compounds such as those  
5 identified in the foregoing assay systems, may be tested for the ability to ameliorate cellular growth or proliferation disorder symptoms. Cell-based and animal model-based assays for the identification of compounds exhibiting such an ability to ameliorate cellular growth or proliferation disorder systems are described herein.

In one aspect, cell-based systems, as described herein, may be used to identify  
10 compounds which may act to ameliorate cellular growth or proliferation disorder symptoms, for example, reduction in tumor burden, tumor size, tumor cell growth, differentiation, and/or proliferation, and invasive and/or metastatic potential before and after treatment. For example, such cell systems may be exposed to a compound, suspected of exhibiting an ability to ameliorate cellular growth or proliferation disorder  
15 symptoms, at a sufficient concentration and for a time sufficient to elicit such an amelioration of cellular growth or proliferation disorder symptoms in the exposed cells. After exposure, the cells are examined to determine whether one or more of the cellular growth or proliferation disorder cellular phenotypes has been altered to resemble a more normal or more wild type, non- cellular growth or proliferation disorder phenotype.  
20 Cellular phenotypes that are associated with cellular growth and/or proliferation disorders include aberrant proliferation, growth, and migration, anchorage independent growth, and loss of contact inhibition.

In addition, animal-based cellular growth or proliferation disorder systems, such as those described herein, may be used to identify compounds capable of ameliorating  
25 cellular growth or proliferation disorder symptoms. Such animal models may be used as test substrates for the identification of drugs, pharmaceuticals, therapies, and interventions which may be effective in treating cellular growth or proliferation disorders. For example, animal models may be exposed to a compound suspected of exhibiting an ability to ameliorate cellular growth or proliferation disorder symptoms, at  
30 a sufficient concentration and for a time sufficient to elicit such an amelioration of cellular growth or proliferation disorder symptoms in the exposed animals. The response of the animals to the exposure may be monitored by assessing the reversal of cellular

growth or proliferation disorders, or symptoms associated therewith, for example, reduction in tumor burden, tumor size, and invasive and/or metastatic potential before and after treatment.

With regard to intervention, any treatments which reverse any aspect of cellular growth or proliferation disorder symptoms should be considered as candidates for human cellular growth or proliferation disorder therapeutic intervention. Dosages of test compounds may be determined by deriving dose-response curves.

Additionally, gene expression patterns may be utilized to assess the ability of a compound to ameliorate cellular growth and/or proliferation disorder symptoms. For example, the expression pattern of one or more genes may form part of a "gene expression profile" or "transcriptional profile" which may be then be used in such an assessment. "Gene expression profile" or "transcriptional profile", as used herein, includes the pattern of mRNA expression obtained for a given tissue or cell type under a given set of conditions. Such conditions may include, but are not limited to, cell growth, proliferation, differentiation, transformation, tumorigenesis, metastasis, and carcinogen exposure. Gene expression profiles may be generated, for example, by utilizing a differential display procedure, Northern analysis and/or RT-PCR. In one embodiment, VR-3 or VR-5 gene sequences may be used as probes and/or PCR primers for the generation and corroboration of such gene expression profiles.

Gene expression profiles may be characterized for known states within the cell- and/or animal-based model systems. Subsequently, these known gene expression profiles may be compared to ascertain the effect a test compound has to modify such gene expression profiles, and to cause the profile to more closely resemble that of a more desirable profile.

For example, administration of a compound may cause the gene expression profile of a cellular growth or proliferation disorder model system to more closely resemble the control system. Administration of a compound may, alternatively, cause the gene expression profile of a control system to begin to mimic a cellular growth and/or proliferation disorder state. Such a compound may, for example, be used in further characterizing the compound of interest, or may be used in the generation of additional animal models.

## B. Cell- and Animal-Based Model Systems

Described herein are cell- and animal-based systems which act as models for cellular growth or proliferation disorders. These systems may be used in a variety of applications. For example, the cell- and animal-based model systems may be used to further characterize differentially expressed genes associated with cellular growth or proliferation disorder, *e.g.*, VR-3 or VR-5. In addition, animal- and cell-based assays may be used as part of screening strategies designed to identify compounds which are capable of ameliorating cellular growth or proliferation disorder symptoms, as described, below. Thus, the animal- and cell-based models may be used to identify drugs, pharmaceuticals, therapies and interventions which may be effective in treating cellular growth or proliferation disorders.

### 1. Animal-Based Systems

Animal-based model systems of cellular growth or proliferation disorders may include, but are not limited to, non-recombinant and engineered transgenic animals.

Animal based models for studying tumorigenesis *in vivo* are well known in the art (reviewed in Animal Models of Cancer Predisposition Syndromes, Hiai, H and Hino, O (eds.) 1999, *Progress in Experimental Tumor Research*, Vol. 35; Clarke AR *Carcinogenesis* (2000) 21:435-41) and include, for example, carcinogen-induced tumors (Rithidech, K *et al. Mutat Res* (1999) 428:33-39; Miller, ML *et al. Environ Mol Mutagen* (2000) 35:319-327), injection and/or transplantation of tumor cells into an animal, as well as animals bearing mutations in growth regulatory genes, for example, oncogenes (*e.g.*, *ras*) (Arbeit, JM *et al. Am J Pathol* (1993) 142:1187-1197; Sinn, E *et al. Cell* (1987) 49:465-475; Thorgeirsson, SS *et al. Toxicol Lett* (2000) 112-113:553-555) and tumor suppressor genes (*e.g.*, *p53*) (Vooijs, M *et al. Oncogene* (1999) 18:5293-5303; Clark AR *Cancer Metast Rev* (1995) 14:125-148; Kumar, TR *et al. J Intern Med* (1995) 238:233-238; Donehower, LA *et al.* (1992) *Nature* 356:215-221). Furthermore, experimental model systems are available for the study of, for example, colon cancer (Heyer J, *et al.* (1999) *Oncogene* 18(38):5325-33), ovarian cancer (Hamilton, TC *et al. Semin Oncol* (1984) 11:285-298; Rahman, NA *et al. Mol Cell Endocrinol* (1998) 145:167-174; Beamer, WG *et al. Toxicol Pathol* (1998) 26:704-710), gastric cancer (Thompson, J *et al. Int J Cancer* (2000) 86:863-869; Fodde, R *et al. Cytogenet Cell*

*Genet* (1999) 86:105-111), breast cancer (Li, M *et al. Oncogene* (2000) 19:1010-1019; Green, JE *et al. Oncogene* (2000) 19:1020-1027), melanoma (Satyamoorthy, K *et al. Cancer Metast Rev* (1999) 18:401-405), and prostate cancer (Shirai, T *et al. Mutat Res* (2000) 462:219-226; Bostwick, DG *et al. Prostate* (2000) 43:286-294).

5           Additionally, animal models exhibiting cellular growth or proliferation disorder symptoms may be engineered by using, for example, VR-3 or VR-5 gene sequences described above, in conjunction with techniques for producing transgenic animals that are well known to those of skill in the art. For example, VR-3 or VR-5 gene sequences may be introduced into, and overexpressed in, the genome of the animal of interest, or, if  
10   endogenous VR-3 or VR-5 gene sequences are present, they may either be overexpressed or, alternatively, be disrupted in order to underexpress or inactivate VR-3 or VR-5 gene expression, such as described for the disruption of apoE in mice (Plump *et al.*, 1992, *Cell* 71: 343-353).

          The host cells of the invention can also be used to produce non-human transgenic  
15   animals. For example, in one embodiment, a host cell of the invention is a fertilized oocyte or an embryonic stem cell into which VR-3 or VR-5-coding sequences have been introduced. Such host cells can then be used to create non-human transgenic animals in which exogenous VR-3 or VR-5 sequences have been introduced into their genome or homologous recombinant animals in which endogenous VR-3 or VR-5 sequences have  
20   been altered. Such animals are useful for studying the function and/or activity of a VR-3 or VR-5 and for identifying and/or evaluating modulators of VR-3 or VR-5 activity. As used herein, a "transgenic animal" is a non-human animal, preferably a mammal, more preferably a rodent such as a rat or mouse, in which one or more of the cells of the animal includes a transgene. Other examples of transgenic animals include non-human  
25   primates, sheep, dogs, cows, goats, chickens, amphibians, and the like. A transgene is exogenous DNA which is integrated into the genome of a cell from which a transgenic animal develops and which remains in the genome of the mature animal, thereby directing the expression of an encoded gene product in one or more cell types or tissues of the transgenic animal. As used herein, a "homologous recombinant animal" is a non-  
30   human animal, preferably a mammal, more preferably a mouse, in which an endogenous VR-3 or VR-5 gene has been altered by homologous recombination between the

endogenous gene and an exogenous DNA molecule introduced into a cell of the animal, *e.g.*, an embryonic cell of the animal, prior to development of the animal.

A transgenic animal of the invention can be created using the methods described herein. VR-3 or VR-5 transgenic animals that express VR-3 or VR-5 mRNA or a VR-3  
5 or VR-5 peptide (detected immunocytochemically, using antibodies directed against VR-3 or VR-5 epitopes) at easily detectable levels should then be further evaluated to identify those animals which display characteristic cellular growth or proliferation disorder symptoms. Tumorigenic disease symptoms include, for example, tumor burden, invasion and/or metastasis.

10 Additionally, specific cell types (*e.g.*, tumor cells, prostate cells, colon cells, breast cells, lung cells, or ovarian cells) within the transgenic animals may be analyzed and assayed for cellular phenotypes characteristic of cellular growth or proliferation disorders. In the case of endothelial cells, such phenotypes include, but are not limited to cell proliferation, growth and migration. Cellular phenotypes associated with a  
15 tumorigenic disorder include, for example, dysregulated proliferation and migration, anchorage independent growth, and loss of contact inhibition. Cellular phenotypes may include a particular cell type's pattern of expression of genes associated with cellular growth or proliferation disorders as compared to known expression profiles of the particular cell type in animals exhibiting cellular growth or proliferation disorder  
20 symptoms.

## 2. Cell-Based Systems

Cells that contain and express VR-3 or VR-5 gene sequences which encode a VR-3 or VR-5 protein, and, further, exhibit cellular phenotypes associated with cellular  
25 growth or proliferation disorders, may be used to identify compounds that exhibit anti-tumorigenic disease activity. Such cells may include endothelial cells such as human umbilical vein endothelial cells (HUVECs), human microvascular endothelial cells (HMVEC); tumor cell lines such as HT-1080 (ATCC# CCL-121), HCT-15 (ATCC# CCL-225), HCC70 (ATCC# CRL-2315), M059J (ATCC# CRL-2366), and NCI-N417  
30 (ATCC# CRL-5809); as well as generic mammalian cell lines such as HeLa cells and COS cells, *e.g.*, COS-7 (ATCC# CRL-1651). Further, such cells may include recombinant, transgenic cell lines. For example, the cellular growth or proliferation

disorder animal models of the invention, discussed above, may be used to generate cell lines, containing one or more cell types involved in cellular growth or proliferation disorders, that can be used as cell culture models for this disorder. While primary cultures derived from the cellular growth or proliferation disorder transgenic animals of the invention may be utilized, the generation of continuous cell lines is preferred. For  
5 examples of techniques which may be used to derive a continuous cell line from the transgenic animals, see Small *et al.*, (1985) *Mol. Cell Biol.* 5:642-648.

Alternatively, cells of a cell type known to be involved in cellular growth or proliferation disorders may be transfected with sequences capable of increasing or  
10 decreasing the amount of VR-3 or VR-5 gene expression within the cell. For example, VR-3 or VR-5 gene sequences may be introduced into, and overexpressed in, the genome of the cell of interest, or, if endogenous VR-3 or VR-5 gene sequences are present, they may be either overexpressed or, alternatively disrupted in order to underexpress or inactivate VR-3 or VR-5 gene expression.

15 In order to overexpress an VR-3 or VR-5 gene, the coding portion of the VR-3 or VR-5 gene may be ligated to a regulatory sequence which is capable of driving gene expression in the cell type of interest, *e.g.*, a tumor cell or a colon cell, a prostate cell, a breast cell, a lung cell, or an ovarian cell. Such regulatory regions will be well known to those of skill in the art, and may be utilized in the absence of undue experimentation.  
20 Recombinant methods for expressing target genes are described above.

For underexpression of an endogenous VR-3 or VR-5 gene sequence, such a sequence may be isolated and engineered such that when reintroduced into the genome of the cell type of interest, the endogenous VR-3 or VR-5 alleles will be inactivated. Preferably, the engineered VR-3 or VR-5 sequence is introduced via gene targeting such  
25 that the endogenous VR-3 or VR-5 sequence is disrupted upon integration of the engineered VR-3 or VR-5 sequence into the cell's genome. Transfection of host cells with VR-3 or VR-5 genes is discussed, above.

Cells treated with compounds or transfected with VR-3 or VR-5 genes can be examined for phenotypes associated with cellular growth or proliferation disorders.

30 Cells (*e.g.*, tumor cells) can be treated with test compounds or transfected with genetically engineered VR-3 or VR-5 genes and examined for phenotypes associated with tumorigenic disease, including, but not limited to changes in cellular morphology,



cell proliferation, cell migration, cell transformation, anchorage independent growth, and loss of contact inhibition.

Transfection of VR-3 or VR-5 nucleic acid may be accomplished by using standard techniques (described in, for example, Ausubel (1989) *supra*). Transfected  
5 cells should be evaluated for the presence of the recombinant VR-3 or VR-5 gene sequences, for expression and accumulation of VR-3 or VR-5 mRNA, and for the presence of recombinant VR-3 or VR-5 protein production. In instances wherein a decrease in VR-3 or VR-5 gene expression is desired, standard techniques may be used to demonstrate whether a decrease in endogenous VR-3 or VR-5 gene expression and/or  
10 in VR-3 or VR-5 protein production is achieved.

Cellular models for the study of tumorigenesis are known in the art, and include cell lines derived from clinical tumors, cells exposed to chemotherapeutic agents, cells exposed to carcinogenic agents, and cell lines with genetic alterations in growth regulatory genes, for example, oncogenes (*e.g.*, *ras*) and tumor suppressor genes (*e.g.*,  
15 *p53*).

This invention further pertains to novel agents identified by the above-described screening assays. Accordingly, it is within the scope of this invention to further use an agent identified as described herein in an appropriate animal model. For example, an agent identified as described herein (*e.g.*, a VR-3 or VR-5 modulating agent, an  
20 antisense VR-3 or VR-5 nucleic acid molecule, a VR-3- or VR5-specific antibody, or a VR-3- or VR5-binding partner) can be used in an animal model to determine the efficacy, toxicity, or side effects of treatment with such an agent. Alternatively, an agent identified as described herein can be used in an animal model to determine the mechanism of action of such an agent. Furthermore, this invention pertains to uses of  
25 novel agents identified by the above-described screening assays for treatments as described herein.

### C. Detection Assays

Portions or fragments of the cDNA sequences identified herein (and the  
30 corresponding complete gene sequences) can be used in numerous ways as polynucleotide reagents. For example, these sequences can be used to: (i) map their respective genes on a chromosome; and, thus, locate gene regions associated with

genetic disease; (ii) identify an individual from a minute biological sample (tissue typing); and (iii) aid in forensic identification of a biological sample. These applications are described in the subsections below.

5           1. Chromosome Mapping

Once the sequence (or a portion of the sequence) of a gene has been isolated, this sequence can be used to map the location of the gene on a chromosome. This process is called chromosome mapping. Accordingly, portions or fragments of the VR-3 or VR-5 nucleotide sequences, described herein, can be used to map the location of the VR-3 or  
10 VR-5 genes on a chromosome. The mapping of the VR-3 or VR-5 sequences to chromosomes is an important first step in correlating these sequences with genes associated with disease.

Briefly, VR-3 or VR-5 genes can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp in length) from the VR-3 or VR-5 nucleotide sequences.  
15 Computer analysis of the VR-3 or VR-5 sequences can be used to predict primers that do not span more than one exon in the genomic DNA, thus complicating the amplification process. These primers can then be used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the VR-3 or VR-5 sequences will yield an amplified  
20 fragment.

Somatic cell hybrids are prepared by fusing somatic cells from different mammals (*e.g.*, human and mouse cells). As hybrids of human and mouse cells grow and divide, they gradually lose human chromosomes in random order, but retain the mouse chromosomes. By using media in which mouse cells cannot grow, because they  
25 lack a particular enzyme, but human cells can, the one human chromosome that contains the gene encoding the needed enzyme, will be retained. By using various media, panels of hybrid cell lines can be established. Each cell line in a panel contains either a single human chromosome or a small number of human chromosomes, and a full set of mouse chromosomes, allowing easy mapping of individual genes to specific human  
30 chromosomes. (D'Eustachio P. *et al.* (1983) *Science* 220:919-924). Somatic cell hybrids containing only fragments of human chromosomes can also be produced by using human chromosomes with translocations and deletions.

PCR mapping of somatic cell hybrids is a rapid procedure for assigning a particular sequence to a particular chromosome. Three or more sequences can be assigned per day using a single thermal cycler. Using the VR-3 or VR-5 nucleotide sequences to design oligonucleotide primers, sublocalization can be achieved with panels of fragments from specific chromosomes. Other mapping strategies which can similarly be used to map a VR-3 or VR-5 sequence to its chromosome include *in situ* hybridization (described in Fan, Y. *et al.* (1990) *Proc. Natl. Acad. Sci. USA*, 87:6223-27), pre-screening with labeled flow-sorted chromosomes, and pre-selection by hybridization to chromosome specific cDNA libraries.

Fluorescence *in situ* hybridization (FISH) of a DNA sequence to a metaphase chromosomal spread can further be used to provide a precise chromosomal location in one step. Chromosome spreads can be made using cells whose division has been blocked in metaphase by a chemical such as colcemid that disrupts the mitotic spindle. The chromosomes can be treated briefly with trypsin, and then stained with Giemsa. A pattern of light and dark bands develops on each chromosome, so that the chromosomes can be identified individually. The FISH technique can be used with a DNA sequence as short as 500 or 600 bases. However, clones larger than 1,000 bases have a higher likelihood of binding to a unique chromosomal location with sufficient signal intensity for simple detection. Preferably 1,000 bases, and more preferably 2,000 bases will suffice to get good results at a reasonable amount of time. For a review of this technique, see Verma *et al.*, Human Chromosomes: A Manual of Basic Techniques (Pergamon Press, New York 1988).

Reagents for chromosome mapping can be used individually to mark a single chromosome or a single site on that chromosome, or panels of reagents can be used for marking multiple sites and/or multiple chromosomes. Reagents corresponding to noncoding regions of the genes actually are preferred for mapping purposes. Coding sequences are more likely to be conserved within gene families, thus increasing the chance of cross hybridizations during chromosomal mapping.

Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. (Such data are found, for example, in V. McKusick, Mendelian Inheritance in Man, available on-line through Johns Hopkins University Welch Medical Library).

The relationship between a gene and a disease, mapped to the same chromosomal region, can then be identified through linkage analysis (co-inheritance of physically adjacent genes), described in, for example, Egeland, J. *et al.* (1987) *Nature*, 325:783-787.

5           Moreover, differences in the DNA sequences between individuals affected and unaffected with a disease associated with the VR-3 or VR-5 gene, can be determined. If a mutation is observed in some or all of the affected individuals but not in any unaffected individuals, then the mutation is likely to be the causative agent of the particular disease. Comparison of affected and unaffected individuals generally  
10 involves first looking for structural alterations in the chromosomes, such as deletions or translocations that are visible from chromosome spreads or detectable using PCR based on that DNA sequence. Ultimately, complete sequencing of genes from several individuals can be performed to confirm the presence of a mutation and to distinguish mutations from polymorphisms.

15

## 2. Tissue Typing

The VR-3 or VR-5 sequences of the present invention can also be used to identify individuals from minute biological samples. The United States military, for example, is considering the use of restriction fragment length polymorphism (RFLP) for  
20 identification of its personnel. In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identification. This method does not suffer from the current limitations of "Dog Tags" which can be lost, switched, or stolen, making positive identification difficult. The sequences of the present invention are useful as additional DNA markers  
25 for RFLP (described in U.S. Patent 5,272,057).

Furthermore, the sequences of the present invention can be used to provide an alternative technique which determines the actual base-by-base DNA sequence of selected portions of an individual's genome. Thus, the VR-3 or VR-5 nucleotide sequences described herein can be used to prepare two PCR primers from the 5' and 3'  
30 ends of the sequences. These primers can then be used to amplify an individual's DNA and subsequently sequence it.

- 87 -

Panels of corresponding DNA sequences from individuals, prepared in this manner, can provide unique individual identifications, as each individual will have a unique set of such DNA sequences due to allelic differences. The sequences of the present invention can be used to obtain such identification sequences from individuals and from tissue. The VR-3 or VR-5 nucleotide sequences of the invention uniquely represent portions of the human genome. Allelic variation occurs to some degree in the coding regions of these sequences, and to a greater degree in the noncoding regions. It is estimated that allelic variation between individual humans occurs with a frequency of about once per each 500 bases. Each of the sequences described herein can, to some degree, be used as a standard against which DNA from an individual can be compared for identification purposes. Because greater numbers of polymorphisms occur in the noncoding regions, fewer sequences are necessary to differentiate individuals. The noncoding sequences of SEQ ID NO:1 or 4 can comfortably provide positive individual identification with a panel of perhaps 10 to 1,000 primers which each yield a noncoding amplified sequence of 75-100 bases. If predicted coding sequences, such as those in SEQ ID NO:3 or 6 are used, a more appropriate number of primers for positive individual identification would be 500-2,000.

If a panel of reagents from VR-3 or VR-5 nucleotide sequences described herein is used to generate a unique identification database for an individual, those same reagents can later be used to identify tissue from that individual. Using the unique identification database, positive identification of the individual, living or dead, can be made from extremely small tissue samples.

### 3. Use of Partial VR-3 or VR-5 Sequences in Forensic Biology

DNA-based identification techniques can also be used in forensic biology. Forensic biology is a scientific field employing genetic typing of biological evidence found at a crime scene as a means for positively identifying, for example, a perpetrator of a crime. To make such an identification, PCR technology can be used to amplify DNA sequences taken from very small biological samples such as tissues, *e.g.*, hair or skin, or body fluids, *e.g.*, blood, saliva, or semen found at a crime scene. The amplified sequence can then be compared to a standard, thereby allowing identification of the origin of the biological sample.

- 88 -

The sequences of the present invention can be used to provide polynucleotide reagents, *e.g.*, PCR primers, targeted to specific loci in the human genome, which can enhance the reliability of DNA-based forensic identifications by, for example, providing another "identification marker" (*i.e.* another DNA sequence that is unique to a particular individual). As mentioned above, actual base sequence information can be used for identification as an accurate alternative to patterns formed by restriction enzyme generated fragments. Sequences targeted to noncoding regions of SEQ ID NO:1 or 4 are particularly appropriate for this use as greater numbers of polymorphisms occur in the noncoding regions, making it easier to differentiate individuals using this technique.

10 Examples of polynucleotide reagents include the VR-3 or VR-5 nucleotide sequences or portions thereof, *e.g.*, fragments derived from the noncoding regions of SEQ ID NO:1 or 4, having a length of at least 20 bases, preferably at least 30 bases.

The VR-3 or VR-5 nucleotide sequences described herein can further be used to provide polynucleotide reagents, *e.g.*, labeled or labelable probes which can be used in, for example, an *in situ* hybridization technique, to identify a specific tissue, *e.g.*, brain tissue. This can be very useful in cases where a forensic pathologist is presented with a tissue of unknown origin. Panels of such VR-3 or VR-5 probes can be used to identify tissue by species and/or by organ type.

15

In a similar fashion, these reagents, *e.g.*, VR-3 or VR-5 primers or probes can be used to screen tissue culture for contamination (*i.e.* screen for the presence of a mixture of different types of cells in a culture).

20

#### D. Predictive Medicine:

The present invention also pertains to the field of predictive medicine in which diagnostic assays, prognostic assays, and monitoring clinical trials are used for prognostic (predictive) purposes to thereby treat an individual prophylactically. Accordingly, one aspect of the present invention relates to diagnostic assays for determining VR-3 or VR-5 protein and/or nucleic acid expression as well as VR-3 or VR-5 activity, in the context of a biological sample (*e.g.*, blood, serum, cells, tissue) to thereby determine whether an individual is afflicted with a disease or disorder, or is at risk of developing a disorder, associated with aberrant or unwanted VR-3 or VR-5 expression or activity. The invention also provides for prognostic (or predictive) assays

25

30

for determining whether an individual is at risk of developing a disorder associated with VR-3 or VR-5 protein, nucleic acid expression or activity. For example, mutations in a VR-3 or VR-5 gene can be assayed in a biological sample. Such assays can be used for prognostic or predictive purpose to thereby prophylactically treat an individual prior to the onset of a disorder characterized by or associated with VR-3 or VR-5 protein, nucleic acid expression or activity.

Another aspect of the invention pertains to monitoring the influence of agents (*e.g.*, drugs, compounds) on the expression or activity of VR-3 or VR-5 in clinical trials.

These and other agents are described in further detail in the following sections.

10

#### 1. Diagnostic Assays

An exemplary method for detecting the presence or absence of VR-3 or VR-5 protein or nucleic acid in a biological sample involves obtaining a biological sample from a test subject and contacting the biological sample with a compound or an agent capable of detecting VR-3 or VR-5 protein or nucleic acid (*e.g.*, mRNA, or genomic DNA) that encodes VR-3 or VR-5 protein such that the presence of VR-3 or VR-5 protein or nucleic acid is detected in the biological sample. A preferred agent for detecting VR-3 or VR-5 mRNA or genomic DNA is a labeled nucleic acid probe capable of hybridizing to VR-3 or VR-5 mRNA or genomic DNA. The nucleic acid probe can be, for example, the VR-3 or VR-5 nucleic acid set forth in SEQ ID NO:1,3, 4, or 6, or the DNA insert of the plasmid deposited with ATCC as Accession Number PTA-2013, or a portion thereof, such as an oligonucleotide of at least 15, 30, 50, 100, 250 or 500 nucleotides in length and sufficient to specifically hybridize under stringent conditions to VR-3 or VR-5 mRNA or genomic DNA. Other suitable probes for use in the diagnostic assays of the invention are described herein.

A preferred agent for detecting VR-3 or VR-5 protein is an antibody capable of binding to VR-3 or VR-5 protein, preferably an antibody with a detectable label. Antibodies can be polyclonal, or more preferably, monoclonal. An intact antibody, or a fragment thereof (*e.g.*, Fab or F(ab')<sub>2</sub>) can be used. The term "labeled", with regard to the probe or antibody, is intended to encompass direct labeling of the probe or antibody by coupling (*i.e.*, physically linking) a detectable substance to the probe or antibody, as well as indirect labeling of the probe or antibody by reactivity with another reagent that

30

is directly labeled. Examples of indirect labeling include detection of a primary antibody using a fluorescently labeled secondary antibody and end-labeling of a DNA probe with biotin such that it can be detected with fluorescently labeled streptavidin. The term "biological sample" is intended to include tissues, cells and biological fluids isolated from a subject, as well as tissues, cells and fluids present within a subject. That is, the detection method of the invention can be used to detect VR-3 or VR-5 mRNA, protein, or genomic DNA in a biological sample *in vitro* as well as *in vivo*. For example, *in vitro* techniques for detection of VR-3 or VR-5 mRNA include Northern hybridizations and *in situ* hybridizations. *In vitro* techniques for detection of VR-3 or VR-5 protein include enzyme linked immunosorbent assays (ELISAs), Western blots, immunoprecipitations and immunofluorescence. *In vitro* techniques for detection of VR-3 or VR-5 genomic DNA include Southern hybridizations. Furthermore, *in vivo* techniques for detection of VR-3 or VR-5 protein include introducing into a subject a labeled anti-VR-3 or anti-VR-5 antibody. For example, the antibody can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques.

In one embodiment, the biological sample contains protein molecules from the test subject. Alternatively, the biological sample can contain mRNA molecules from the test subject or genomic DNA molecules from the test subject. A preferred biological sample is a serum sample isolated by conventional means from a subject.

In another embodiment, the methods further involve obtaining a control biological sample from a control subject, contacting the control sample with a compound or agent capable of detecting VR-3 or VR-5 protein, mRNA, or genomic DNA, such that the presence of VR-3 or VR-5 protein, mRNA or genomic DNA is detected in the biological sample, and comparing the presence of VR-3 or VR-5 protein, mRNA or genomic DNA in the control sample with the presence of VR-3 or VR-5 protein, mRNA or genomic DNA in the test sample.

The invention also encompasses kits for detecting the presence of VR-3 or VR-5 in a biological sample. For example, the kit can comprise a labeled compound or agent capable of detecting VR-3 or VR-5 protein or mRNA in a biological sample; means for determining the amount of VR-3 or VR-5 in the sample; and means for comparing the amount of VR-3 or VR-5 in the sample with a standard. The compound or agent can be



packaged in a suitable container. The kit can further comprise instructions for using the kit to detect VR-3 or VR-5 protein or nucleic acid.

## 2. Prognostic Assays

5           The diagnostic methods described herein can furthermore be utilized to identify subjects having or at risk of developing a disease or disorder associated with aberrant or unwanted VR-3 or VR-5 expression or activity. As used herein, the term "aberrant" includes a VR-3 or VR-5 expression or activity which deviates from the wild type VR-3 or VR-5 expression or activity. Aberrant expression or activity includes increased or  
10       decreased expression or activity, as well as expression or activity which does not follow the wild type developmental pattern of expression or the subcellular pattern of expression. For example, aberrant VR-3 or VR-5 expression or activity is intended to include the cases in which a mutation in the VR-3 or VR-5 gene causes the VR-3 or VR-5 gene to be under-expressed or over-expressed and situations in which such mutations  
15       result in a non-functional VR-3 or VR-5 protein or a protein which does not function in a wild-type fashion, *e.g.*, a protein which does not interact with a VR-3 or VR-5 ligand, *e.g.*, a vanilloid compound, or one which interacts with a non-VR-3 or non-VR-5 ligand, *e.g.* a non-vanilloid compound. As used herein, the term "unwanted" includes an unwanted phenomenon involved in a biological response such as aberrant  
20       transmission of pain stimuli, aberrant transport of calcium, or aberrant cellular growth and/or proliferation, or tumorigenesis. For example, the term unwanted includes a VR-3 or VR-5 expression or activity which is undesirable in a subject.

          The assays described herein, such as the preceding diagnostic assays or the following assays, can be utilized to identify a subject having or at risk of developing a  
25       disorder associated with a misregulation in VR-3 or VR-5 protein activity or nucleic acid expression, such as calcium homeostasis related disorders, cellular growth and/or proliferation disorders, *e.g.* cancer, and/or pain disorders. Alternatively, the prognostic assays can be utilized to identify a subject having or at risk for developing a disorder associated with a misregulation in VR-3 or VR-5 protein activity or nucleic acid  
30       expression, such as calcium homeostasis related disorders, cellular growth and/or proliferation disorders, *e.g.* cancer, and/or pain disorders. Thus, the present invention provides a method for identifying a disease or disorder associated with aberrant or

unwanted VR-3 or VR-5 expression or activity in which a test sample is obtained from a subject and VR-3 or VR-5 protein or nucleic acid (*e.g.*, mRNA or genomic DNA) is detected, wherein the presence of VR-3 or VR-5 protein or nucleic acid is diagnostic for a subject having or at risk of developing a disease or disorder associated with aberrant or unwanted VR-3 or VR-5 expression or activity. As used herein, a "test sample" refers to a biological sample obtained from a subject of interest. For example, a test sample can be a biological fluid (*e.g.*, serum), cell sample, or tissue.

Furthermore, the prognostic assays described herein can be used to determine whether a subject can be administered an agent (*e.g.*, an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate) to treat a disease or disorder associated with aberrant or unwanted VR-3 or VR-5 expression or activity. For example, such methods can be used to determine whether a subject can be effectively treated with an agent for a pain disorder, a calcium homeostasis related disorder, or a cellular growth and/or proliferation disorder. Thus, the present invention provides methods for determining whether a subject can be effectively treated with an agent for a disorder associated with aberrant or unwanted VR-3 or VR-5 expression or activity in which a test sample is obtained and VR-3 or VR-5 protein or nucleic acid expression or activity is detected (*e.g.*, wherein the abundance of VR-3 or VR-5 protein or nucleic acid expression or activity is diagnostic for a subject that can be administered the agent to treat a disorder associated with aberrant or unwanted VR-3 or VR-5 expression or activity).

The methods of the invention can also be used to detect genetic alterations in a VR-3 or VR-5 gene, thereby determining if a subject with the altered gene is at risk for a disorder characterized by misregulation in VR-3 or VR-5 protein activity or nucleic acid expression, such as calcium homeostasis related disorders, cellular growth and/or proliferation disorders, *e.g.* cancer, and/or pain disorders. In preferred embodiments, the methods include detecting, in a sample of cells from the subject, the presence or absence of a genetic alteration characterized by at least one of an alteration affecting the integrity of a gene encoding a VR-3- or VR5-protein, or the mis-expression of the VR-3 or VR-5 gene. For example, such genetic alterations can be detected by ascertaining the existence of at least one of 1) a deletion of one or more nucleotides from a VR-3 or VR-5 gene; 2) an addition of one or more nucleotides to a VR-3 or VR-5 gene; 3) a

substitution of one or more nucleotides of a VR-3 or VR-5 gene, 4) a chromosomal rearrangement of a VR-3 or VR-5 gene; 5) an alteration in the level of a messenger RNA transcript of a VR-3 or VR-5 gene, 6) aberrant modification of a VR-3 or VR-5 gene, such as of the methylation pattern of the genomic DNA, 7) the presence of a non-wild type splicing pattern of a messenger RNA transcript of a VR-3 or VR-5 gene, 8) a non-wild type level of a VR-3- or VR5-protein, 9) allelic loss of a VR-3 or VR-5 gene, and 10) inappropriate post-translational modification of a VR-3- or VR5-protein. As described herein, there are a large number of assays known in the art which can be used for detecting alterations in a VR-3 or VR-5 gene. A preferred biological sample is a tissue or serum sample isolated by conventional means from a subject.

In certain embodiments, detection of the alteration involves the use of a probe/primer in a polymerase chain reaction (PCR) (see, *e.g.*, U.S. Patent Nos. 4,683,195 and 4,683,202), such as anchor PCR or RACE PCR, or, alternatively, in a ligation chain reaction (LCR) (see, *e.g.*, Landegran *et al.* (1988) *Science* 241:1077-1080; and Nakazawa *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:360-364), the latter of which can be particularly useful for detecting point mutations in the VR-3- or VR5-gene (see Abravaya *et al.* (1995) *Nucleic Acids Res.* 23:675-682). This method can include the steps of collecting a sample of cells from a subject, isolating nucleic acid (*e.g.*, genomic, mRNA or both) from the cells of the sample, contacting the nucleic acid sample with one or more primers which specifically hybridize to a VR-3 or VR-5 gene under conditions such that hybridization and amplification of the VR-3- or VR5-gene (if present) occurs, and detecting the presence or absence of an amplification product, or detecting the size of the amplification product and comparing the length to a control sample. It is anticipated that PCR and/or LCR may be desirable to use as a preliminary amplification step in conjunction with any of the techniques used for detecting mutations described herein.

Alternative amplification methods include: self sustained sequence replication (Guatelli, J.C. *et al.*, (1990) *Proc. Natl. Acad. Sci. USA* 87:1874-1878), transcriptional amplification system (Kwoh, D.Y. *et al.*, (1989) *Proc. Natl. Acad. Sci. USA* 86:1173-1177), Q-Beta Replicase (Lizardi, P.M. *et al.* (1988) *Bio-Technology* 6:1197), or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These

detection schemes are especially useful for the detection of nucleic acid molecules if such molecules are present in very low numbers.

In an alternative embodiment, mutations in a VR-3 or VR-5 gene from a sample cell can be identified by alterations in restriction enzyme cleavage patterns. For example, sample and control DNA is isolated, amplified (optionally), digested with one or more restriction endonucleases, and fragment length sizes are determined by gel electrophoresis and compared. Differences in fragment length sizes between sample and control DNA indicates mutations in the sample DNA. Moreover, the use of sequence specific ribozymes (see, for example, U.S. Patent No. 5,498,531) can be used to score for the presence of specific mutations by development or loss of a ribozyme cleavage site.

In other embodiments, genetic mutations in VR-3 or VR-5 can be identified by hybridizing a sample and control nucleic acids, *e.g.*, DNA or RNA, to high density arrays containing hundreds or thousands of oligonucleotides probes (Cronin, M.T. *et al.* (1996) *Human Mutation* 7: 244-255; Kozal, M.J. *et al.* (1996) *Nature Medicine* 2: 753-759). For example, genetic mutations in VR-3 or VR-5 can be identified in two dimensional arrays containing light-generated DNA probes as described in Cronin, M.T. *et al. supra*. Briefly, a first hybridization array of probes can be used to scan through long stretches of DNA in a sample and control to identify base changes between the sequences by making linear arrays of sequential overlapping probes. This step allows the identification of point mutations. This step is followed by a second hybridization array that allows the characterization of specific mutations by using smaller, specialized probe arrays complementary to all variants or mutations detected. Each mutation array is composed of parallel probe sets, one complementary to the wild-type gene and the other complementary to the mutant gene.

In yet another embodiment, any of a variety of sequencing reactions known in the art can be used to directly sequence the VR-3 or VR-5 gene and detect mutations by comparing the sequence of the sample VR-3 or VR-5 with the corresponding wild-type (control) sequence. Examples of sequencing reactions include those based on techniques developed by Maxam and Gilbert ((1977) *Proc. Natl. Acad. Sci. USA* 74:560) or Sanger ((1977) *Proc. Natl. Acad. Sci. USA* 74:5463). It is also contemplated that any of a variety of automated sequencing procedures can be utilized when

performing the diagnostic assays ((1995) *Biotechniques* 19:448), including sequencing by mass spectrometry (see, *e.g.*, PCT International Publication No. WO 94/16101; Cohen *et al.* (1996) *Adv. Chromatogr.* 36:127-162; and Griffin *et al.* (1993) *Appl. Biochem. Biotechnol.* 38:147-159).

5           Other methods for detecting mutations in the VR-3 or VR-5 gene include methods in which protection from cleavage agents is used to detect mismatched bases in RNA/RNA or RNA/DNA heteroduplexes (Myers *et al.* (1985) *Science* 230:1242). In general, the art technique of "mismatch cleavage" starts by providing heteroduplexes of formed by hybridizing (labeled) RNA or DNA containing the wild-type VR-3 or VR-5  
10   sequence with potentially mutant RNA or DNA obtained from a tissue sample. The double-stranded duplexes are treated with an agent which cleaves single-stranded regions of the duplex such as which will exist due to basepair mismatches between the control and sample strands. For instance, RNA/DNA duplexes can be treated with RNase and DNA/DNA hybrids treated with S1 nuclease to enzymatically digesting the  
15   mismatched regions. In other embodiments, either DNA/DNA or RNA/DNA duplexes can be treated with hydroxylamine or osmium tetroxide and with piperidine in order to digest mismatched regions. After digestion of the mismatched regions, the resulting material is then separated by size on denaturing polyacrylamide gels to determine the site of mutation. See, for example, Cotton *et al.* (1988) *Proc. Natl Acad Sci USA*  
20   85:4397; Saleeba *et al.* (1992) *Methods Enzymol.* 217:286-295. In a preferred embodiment, the control DNA or RNA can be labeled for detection.

          In still another embodiment, the mismatch cleavage reaction employs one or more proteins that recognize mismatched base pairs in double-stranded DNA (so called "DNA mismatch repair" enzymes) in defined systems for detecting and mapping point  
25   mutations in VR-3 or VR-5 cDNAs obtained from samples of cells. For example, the mutY enzyme of *E. coli* cleaves A at G/A mismatches and the thymidine DNA glycosylase from HeLa cells cleaves T at G/T mismatches (Hsu *et al.* (1994) *Carcinogenesis* 15:1657-1662). According to an exemplary embodiment, a probe based on a VR-3 or VR-5 sequence, *e.g.*, a wild-type VR-3 or VR-5 sequence, is hybridized to  
30   a cDNA or other DNA product from a test cell(s). The duplex is treated with a DNA mismatch repair enzyme, and the cleavage products, if any, can be detected from electrophoresis protocols or the like. See, for example, U.S. Patent No. 5,459,039.

In other embodiments, alterations in electrophoretic mobility will be used to identify mutations in VR-3 or VR-5 genes. For example, single strand conformation polymorphism (SSCP) may be used to detect differences in electrophoretic mobility between mutant and wild type nucleic acids (orita *et al.* (1989) *Proc Natl. Acad. Sci USA*: 86:2766, see also Cotton (1993) *Mutat. Res.* 285:125-144; and Hayashi (1992) *Genet. Anal. Tech. Appl.* 9:73-79). Single-stranded DNA fragments of sample and control VR-3 or VR-5 nucleic acids will be denatured and allowed to renature. The secondary structure of single-stranded nucleic acids varies according to sequence, the resulting alteration in electrophoretic mobility enables the detection of even a single base change. The DNA fragments may be labeled or detected with labeled probes. The sensitivity of the assay may be enhanced by using RNA (rather than DNA), in which the secondary structure is more sensitive to a change in sequence. In a preferred embodiment, the subject method utilizes heteroduplex analysis to separate double stranded heteroduplex molecules on the basis of changes in electrophoretic mobility (Keen *et al.* (1991) *Trends Genet* 7:5).

In yet another embodiment the movement of mutant or wild-type fragments in polyacrylamide gels containing a gradient of denaturant is assayed using denaturing gradient gel electrophoresis (DGGE) (Myers *et al.* (1985) *Nature* 313:495). When DGGE is used as the method of analysis, DNA will be modified to insure that it does not completely denature, for example by adding a GC clamp of approximately 40 bp of high-melting GC-rich DNA by PCR. In a further embodiment, a temperature gradient is used in place of a denaturing gradient to identify differences in the mobility of control and sample DNA (Rosenbaum and Reissner (1987) *Biophys Chem* 265:12753).

Examples of other techniques for detecting point mutations include, but are not limited to, selective oligonucleotide hybridization, selective amplification, or selective primer extension. For example, oligonucleotide primers may be prepared in which the known mutation is placed centrally and then hybridized to target DNA under conditions which permit hybridization only if a perfect match is found (Saiki *et al.* (1986) *Nature* 324:163); Saiki *et al.* (1989) *Proc. Natl Acad. Sci USA* 86:6230). Such allele specific oligonucleotides are hybridized to PCR amplified target DNA or a number of different mutations when the oligonucleotides are attached to the hybridizing membrane and hybridized with labeled target DNA.

Alternatively, allele specific amplification technology which depends on selective PCR amplification may be used in conjunction with the instant invention. Oligonucleotides used as primers for specific amplification may carry the mutation of interest in the center of the molecule (so that amplification depends on differential hybridization) (Gibbs *et al.* (1989) *Nucleic Acids Res.* 17:2437-2448) or at the extreme 3' end of one primer where, under appropriate conditions, mismatch can prevent, or reduce polymerase extension (Prossner (1993) *Tibtech* 11:238). In addition it may be desirable to introduce a novel restriction site in the region of the mutation to create cleavage-based detection (Gasparini *et al.* (1992) *Mol. Cell Probes* 6:1). It is anticipated that in certain embodiments amplification may also be performed using Taq ligase for amplification (Barany (1991) *Proc. Natl. Acad. Sci USA* 88:189). In such cases, ligation will occur only if there is a perfect match at the 3' end of the 5' sequence making it possible to detect the presence of a known mutation at a specific site by looking for the presence or absence of amplification.

The methods described herein may be performed, for example, by utilizing pre-packaged diagnostic kits comprising at least one probe nucleic acid or antibody reagent described herein, which may be conveniently used, *e.g.*, in clinical settings to diagnose patients exhibiting symptoms or family history of a disease or illness involving a VR-3 or VR-5 gene.

Furthermore, any cell type or tissue in which VR-3 or VR-5 is expressed may be utilized in the prognostic assays described herein.

### 3. Monitoring of Effects During Clinical Trials

Monitoring the influence of agents (*e.g.*, drugs) on the expression or activity of a VR-3 or VR-5 protein (*e.g.*, the modulation of pain signaling mechanisms, the regulation of calcium homeostasis, or the modulation of cellular growth and/or proliferation) can be applied not only in basic drug screening, but also in clinical trials. For example, the effectiveness of an agent determined by a screening assay as described herein to increase VR-3 or VR-5 gene expression, protein levels, or upregulate VR-3 or VR-5 activity, can be monitored in clinical trials of subjects exhibiting decreased VR-3 or VR-5 gene expression, protein levels, or downregulated VR-3 or VR-5 activity. Alternatively, the effectiveness of an agent determined by a screening assay to decrease

VR-3 or VR-5 gene expression, protein levels, or downregulate VR-3 or VR-5 activity, can be monitored in clinical trials of subjects exhibiting increased VR-3 or VR-5 gene expression, protein levels, or upregulated VR-3 or VR-5 activity. In such clinical trials, the expression or activity of a VR-3 or VR-5 gene, and preferably, other genes that have  
5 been implicated in, for example, a VR-3- or VR5-associated disorder can be used as a "read out" or markers of the phenotype of a particular cell.

For example, and not by way of limitation, genes, including VR-3 or VR-5, that are modulated in cells by treatment with an agent (*e.g.*, compound, drug or small molecule) which modulates VR-3 or VR-5 activity (*e.g.*, identified in a screening assay  
10 as described herein) can be identified. Thus, to study the effect of agents on VR-3- or VR5-associated disorders (*e.g.*, calcium homeostasis related disorders; pain disorders, and cellular growth and/or proliferation disorders, *e.g.*, cancer), for example, in a clinical trial, cells can be isolated and RNA prepared and analyzed for the levels of expression of VR-3 or VR-5 and other genes implicated in the VR-3- or VR5-associated  
15 disorder, respectively. The levels of gene expression (*e.g.*, a gene expression pattern) can be quantified by northern blot analysis or RT-PCR, as described herein, or alternatively by measuring the amount of protein produced, by one of the methods as described herein, or by measuring the levels of activity of VR-3 or VR-5 or other genes. In this way, the gene expression pattern can serve as a marker, indicative of the  
20 physiological response of the cells to the agent. Accordingly, this response state may be determined before, and at various points during treatment of the individual with the agent.

In a preferred embodiment, the present invention provides a method for monitoring the effectiveness of treatment of a subject with an agent (*e.g.*, an agonist,  
25 antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate identified by the screening assays described herein) including the steps of (i) obtaining a pre-administration sample from a subject prior to administration of the agent; (ii) detecting the level of expression of a VR-3 or VR-5 protein, mRNA, or genomic DNA in the preadministration sample; (iii) obtaining one or more post-  
30 administration samples from the subject; (iv) detecting the level of expression or activity of the VR-3 or VR-5 protein, mRNA, or genomic DNA in the post-administration samples; (v) comparing the level of expression or activity of the VR-3 or VR-5 protein,



mRNA, or genomic DNA in the pre-administration sample with the VR-3 or VR-5 protein, mRNA, or genomic DNA in the post administration sample or samples; and (vi) altering the administration of the agent to the subject accordingly. For example, increased administration of the agent may be desirable to increase the expression or activity of VR-3 or VR-5 to higher levels than detected, *i.e.*, to increase the effectiveness of the agent. Alternatively, decreased administration of the agent may be desirable to decrease expression or activity of VR-3 or VR-5 to lower levels than detected, *i.e.* to decrease the effectiveness of the agent. According to such an embodiment, VR-3 or VR-5 expression or activity may be used as an indicator of the effectiveness of an agent, even in the absence of an observable phenotypic response.

#### E. Methods of Treatment:

The present invention provides for both prophylactic and therapeutic methods of treating a subject at risk of (or susceptible to) a disorder or having a disorder associated with aberrant or unwanted VR-3 or VR-5 expression or activity, *e.g.*, a calcium homeostasis related disorder, pain or a pain disorder, or a cellular growth and/or proliferation disorder, *e.g.*, cancer. "Treatment", or "treating" as used herein, is defined as the application or administration of a therapeutic agent to a patient, or application or administration of a therapeutic agent to an isolated tissue or cell line from a patient, who has a disease or disorder, a symptom of disease or disorder or a predisposition toward a disease or disorder, with the purpose to cure, heal, alleviate, relieve, alter, remedy, ameliorate, improve or affect the disease or disorder, the symptoms of the disease or disorder, or the predisposition toward disease. A therapeutic agent includes, but is not limited to, small molecules, peptides, antibodies, ribozymes and antisense oligonucleotides. With regard to both prophylactic and therapeutic methods of treatment, such treatments may be specifically tailored or modified, based on knowledge obtained from the field of pharmacogenomics. "Pharmacogenomics", as used herein, refers to the application of genomics technologies such as gene sequencing, statistical genetics, and gene expression analysis to drugs in clinical development and on the market. More specifically, the term refers the study of how a patient's genes determine his or her response to a drug (*e.g.*, a patient's "drug response phenotype", or "drug response genotype"). Thus, another aspect of the invention provides methods for

tailoring an individual's prophylactic or therapeutic treatment with either the VR-3 or VR-5 molecules of the present invention or VR-3 or VR-5 modulators according to that individual's drug response genotype. Pharmacogenomics allows a clinician or physician to target prophylactic or therapeutic treatments to patients who will most benefit from the treatment and to avoid treatment of patients who will experience toxic drug-related side effects.

### 1. Prophylactic Methods

In one aspect, the invention provides a method for preventing in a subject, a disease or condition associated with an aberrant or unwanted VR-3 or VR-5 expression or activity, by administering to the subject a VR-3 or VR-5 or an agent which modulates VR-3 or VR-5 expression or at least one VR-3 or VR-5 activity. Subjects at risk for a disease which is caused or contributed to by aberrant or unwanted VR-3 or VR-5 expression or activity can be identified by, for example, any or a combination of diagnostic or prognostic assays as described herein. Administration of a prophylactic agent can occur prior to the manifestation of symptoms characteristic of the VR-3 or VR-5 aberrancy, such that a disease or disorder is prevented or, alternatively, delayed in its progression. Depending on the type of VR-3 or VR-5 aberrancy, for example, a VR-3 or VR-5, VR-3 or VR-5 agonist or VR-3 or VR-5 antagonist agent can be used for treating the subject. The appropriate agent can be determined based on screening assays described herein.

### 2. Therapeutic Methods

Another aspect of the invention pertains to methods of modulating VR-3 or VR-5 expression or activity for therapeutic purposes. Accordingly, in an exemplary embodiment, the modulatory method of the invention involves contacting a cell with a VR-3 or VR-5 or agent that modulates one or more of the activities of VR-3 or VR-5 protein activity associated with the cell. An agent that modulates VR-3 or VR-5 protein activity can be an agent as described herein, such as a nucleic acid or a protein, a naturally-occurring target molecule of a VR-3 or VR-5 protein (*e.g.*, a VR-3 or VR-5 ligand or substrate), a VR-3 or VR-5 antibody, a VR-3 or VR-5 agonist or antagonist, a peptidomimetic of a VR-3 or VR-5 agonist or antagonist, or other small molecule. In

one embodiment, the agent stimulates one or more VR-3 or VR-5 activities. Examples of such stimulatory agents include active VR-3 or VR-5 protein and a nucleic acid molecule encoding VR-3 or VR-5 that has been introduced into the cell. In another embodiment, the agent inhibits one or more VR-3 or VR-5 activities. Examples of such inhibitory agents include antisense VR-3 or VR-5 nucleic acid molecules, anti-VR-3 or anti-VR-5 antibodies, and VR-3 or VR-5 inhibitors. These modulatory methods can be performed *in vitro* (e.g., by culturing the cell with the agent) or, alternatively, *in vivo* (e.g., by administering the agent to a subject). As such, the present invention provides methods of treating an individual afflicted with a disease or disorder characterized by aberrant or unwanted expression or activity of a VR-3 or VR-5 protein or nucleic acid molecule such as a calcium homeostasis related disorder, a pain disorder, or a cellular growth and/or proliferation disorder, e.g., cancer. In one embodiment, the method involves administering an agent (e.g., an agent identified by a screening assay described herein), or combination of agents that modulates (e.g., upregulates or downregulates) VR-3 or VR-5 expression or activity. In another embodiment, the method involves administering a VR-3 or VR-5 protein or nucleic acid molecule as therapy to compensate for reduced, aberrant, or unwanted VR-3 or VR-5 expression or activity.

Stimulation of VR-3 or VR-5 activity is desirable in situations in which VR-3 or VR-5 is abnormally downregulated and/or in which increased VR-3 or VR-5 activity is likely to have a beneficial effect. Likewise, inhibition of VR-3 or VR-5 activity is desirable in situations in which VR-3 or VR-5 is abnormally upregulated and/or in which decreased VR-3 or VR-5 activity is likely to have a beneficial effect.

### 3. Pharmacogenomics

The VR-3 or VR-5 molecules of the present invention, as well as agents, or modulators which have a stimulatory or inhibitory effect on VR-3 or VR-5 activity (e.g., VR-3 or VR-5 gene expression) as identified by a screening assay described herein can be administered to individuals to treat (prophylactically or therapeutically) VR-3- or VR5-associated disorders (e.g., calcium homeostasis related disorders; pain disorders; and cellular growth and/or proliferation disorders, e.g., cancer) associated with aberrant or unwanted VR-3 or VR-5 activity. In conjunction with such treatment, pharmacogenomics (*i.e.*, the study of the relationship between an individual's genotype

and that individual's response to a foreign compound or drug) may be considered. Differences in metabolism of therapeutics can lead to severe toxicity or therapeutic failure by altering the relation between dose and blood concentration of the pharmacologically active drug. Thus, a physician or clinician may consider applying knowledge obtained in relevant pharmacogenomics studies in determining whether to administer a VR-3 or VR-5 molecule or VR-3 or VR-5 modulator as well as tailoring the dosage and/or therapeutic regimen of treatment with a VR-3 or VR-5 molecule or VR-3 or VR-5 modulator.

Pharmacogenomics deals with clinically significant hereditary variations in the response to drugs due to altered drug disposition and abnormal action in affected persons. See, for example, Eichelbaum, M. *et al.* (1996) *Clin. Exp. Pharmacol. Physiol.* 23(10-11): 983-985 and Linder, M.W. *et al.* (1997) *Clin. Chem.* 43(2):254-266. In general, two types of pharmacogenetic conditions can be differentiated. Genetic conditions transmitted as a single factor altering the way drugs act on the body (altered drug action) or genetic conditions transmitted as single factors altering the way the body acts on drugs (altered drug metabolism). These pharmacogenetic conditions can occur either as rare genetic defects or as naturally-occurring polymorphisms. For example, glucose-6-phosphate dehydrogenase deficiency (G6PD) is a common inherited enzymopathy in which the main clinical complication is haemolysis after ingestion of oxidant drugs (anti-malarials, sulfonamides, analgesics, nitrofurans) and consumption of fava beans.

One pharmacogenomics approach to identifying genes that predict drug response, known as "a genome-wide association", relies primarily on a high-resolution map of the human genome consisting of already known gene-related markers (*e.g.*, a "bi-allelic" gene marker map which consists of 60,000-100,000 polymorphic or variable sites on the human genome, each of which has two variants.) Such a high-resolution genetic map can be compared to a map of the genome of each of a statistically significant number of patients taking part in a Phase II/III drug trial to identify markers associated with a particular observed drug response or side effect. Alternatively, such a high resolution map can be generated from a combination of some ten-million known single nucleotide polymorphisms (SNPs) in the human genome. As used herein, a "SNP" is a common alteration that occurs in a single nucleotide base in a stretch of

DNA. For example, a SNP may occur once per every 1000 bases of DNA. A SNP may be involved in a disease process, however, the vast majority may not be disease-associated. Given a genetic map based on the occurrence of such SNPs, individuals can be grouped into genetic categories depending on a particular pattern of SNPs in their individual genome. In such a manner, treatment regimens can be tailored to groups of genetically similar individuals, taking into account traits that may be common among such genetically similar individuals.

Alternatively, a method termed the "candidate gene approach", can be utilized to identify genes that predict drug response. According to this method, if a gene that encodes a drug's target is known (*e.g.*, a VR-3 or VR-5 protein of the present invention), all common variants of that gene can be fairly easily identified in the population and it can be determined if having one version of the gene versus another is associated with a particular drug response.

As an illustrative embodiment, the activity of drug metabolizing enzymes is a major determinant of both the intensity and duration of drug action. The discovery of genetic polymorphisms of drug metabolizing enzymes (*e.g.*, N-acetyltransferase 2 (NAT 2) and cytochrome P450 enzymes CYP2D6 and CYP2C19) has provided an explanation as to why some patients do not obtain the expected drug effects or show exaggerated drug response and serious toxicity after taking the standard and safe dose of a drug. These polymorphisms are expressed in two phenotypes in the population, the extensive metabolizer (EM) and poor metabolizer (PM). The prevalence of PM is different among different populations. For example, the gene coding for CYP2D6 is highly polymorphic and several mutations have been identified in PM, which all lead to the absence of functional CYP2D6. Poor metabolizers of CYP2D6 and CYP2C19 quite frequently experience exaggerated drug response and side effects when they receive standard doses. If a metabolite is the active therapeutic moiety, PM show no therapeutic response, as demonstrated for the analgesic effect of codeine mediated by its CYP2D6-formed metabolite morphine. The other extreme are the so called ultra-rapid metabolizers who do not respond to standard doses. Recently, the molecular basis of ultra-rapid metabolism has been identified to be due to CYP2D6 gene amplification.

Alternatively, a method termed the "gene expression profiling", can be utilized to identify genes that predict drug response. For example, the gene expression of an animal dosed with a drug (*e.g.*, a VR-3 or VR-5 molecule or VR-3 or VR-5 modulator of the present invention) can give an indication whether gene pathways related to toxicity have been turned on.

Information generated from more than one of the above pharmacogenomics approaches can be used to determine appropriate dosage and treatment regimens for prophylactic or therapeutic treatment an individual. This knowledge, when applied to dosing or drug selection, can avoid adverse reactions or therapeutic failure and thus enhance therapeutic or prophylactic efficiency when treating a subject with a VR-3 or VR-5 molecule or VR-3 or VR-5 modulator, such as a modulator identified by one of the exemplary screening assays described herein.

#### VI. Electronic Apparatus Readable Media and Arrays

Electronic apparatus readable media comprising VR-3 or VR-5 sequence information is also provided. As used herein, "VR-3 or VR-5 sequence information" refers to any nucleotide and/or amino acid sequence information particular to the VR-3 or VR-5 molecules of the present invention, including but not limited to full-length nucleotide and/or amino acid sequences, partial nucleotide and/or amino acid sequences, polymorphic sequences including single nucleotide polymorphisms (SNPs), epitope sequences, and the like. Moreover, information "related to" said VR-3 or VR-5 sequence information includes detection of the presence or absence of a sequence (*e.g.*, detection of expression of a sequence, fragment, polymorphism, etc.), determination of the level of a sequence (*e.g.*, detection of a level of expression, for example, a quantitative detection), detection of a reactivity to a sequence (*e.g.*, detection of protein expression and/or levels, for example, using a sequence-specific antibody), and the like. As used herein, "electronic apparatus readable media" refers to any suitable medium for storing, holding or containing data or information that can be read and accessed directly by an electronic apparatus. Such media can include, but are not limited to: magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as compact disc; electronic storage media such as RAM, ROM, EPROM, EEPROM and the like; general hard disks and hybrids of these categories such

as magnetic/optical storage media. The medium is adapted or configured for having recorded thereon VR-3 or VR-5 sequence information of the present invention.

As used herein, the term "electronic apparatus" is intended to include any suitable computing or processing apparatus or other device configured or adapted for storing data or information. Examples of electronic apparatus suitable for use with the present invention include stand-alone computing apparatus; networks, including a local area network (LAN), a wide area network (WAN) Internet, Intranet, and Extranet; electronic appliances such as a personal digital assistants (PDAs), cellular phone, pager and the like; and local and distributed processing systems.

As used herein, "recorded" refers to a process for storing or encoding information on the electronic apparatus readable medium. Those skilled in the art can readily adopt any of the presently known methods for recording information on known media to generate manufactures comprising the VR-3 or VR-5 sequence information.

A variety of software programs and formats can be used to store the sequence information on the electronic apparatus readable medium. For example, the sequence information can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and MicroSoft Word, or represented in the form of an ASCII file, stored in a database application, such as DB2, Sybase, Oracle, or the like, as well as in other forms. Any number of data processor structuring formats (*e.g.*, text file or database) may be employed in order to obtain or create a medium having recorded thereon the VR-3 or VR-5 sequence information.

By providing VR-3 or VR-5 sequence information in readable form, one can routinely access the sequence information for a variety of purposes. For example, one skilled in the art can use the sequence information in readable form to compare a target sequence or target structural motif with the sequence information stored within the data storage means. Search means are used to identify fragments or regions of the sequences of the invention which match a particular target sequence or target motif.

The present invention therefore provides a medium for holding instructions for performing a method for determining whether a subject has a VR-3 or VR-5-associated disease or disorder or a pre-disposition to a VR-3 or VR-5-associated disease or disorder, wherein the method comprises the steps of determining VR-3 or VR-5 sequence information associated with the subject and based on the VR-3 or VR-5

sequence information, determining whether the subject has a VR-3 or VR-5-associated disease or disorder or a pre-disposition to a VR-3 or VR-5-associated disease or disorder and/or recommending a particular treatment for the disease, disorder or pre-disease condition.

5           The present invention further provides in an electronic system and/or in a network, a method for determining whether a subject has a VR-3 or VR-5-associated disease or disorder or a pre-disposition to a disease associated with a VR-3 or VR-5 wherein the method comprises the steps of determining VR-3 or VR-5 sequence information associated with the subject, and based on the VR-3 or VR-5 sequence  
10 information, determining whether the subject has a VR-3 or VR-5-associated disease or disorder or a pre-disposition to a VR-3 or VR-5-associated disease or disorder, and/or recommending a particular treatment for the disease, disorder or pre-disease condition. The method may further comprise the step of receiving phenotypic information associated with the subject and/or acquiring from a network phenotypic information  
15 associated with the subject.

          The present invention also provides in a network, a method for determining whether a subject has a VR-3 or VR-5 -associated disease or disorder or a pre-disposition to a VR-3 or VR-5-associated disease or disorder associated with VR-3 or VR-5, said method comprising the steps of receiving VR-3 or VR-5 sequence  
20 information from the subject and/or information related thereto, receiving phenotypic information associated with the subject, acquiring information from the network corresponding to VR-3 or VR-5 and/or a [VR-3 or VR-5]-associated disease or disorder, and based on one or more of the phenotypic information, the VR-3 or VR-5 information (*e.g.*, sequence information and/or information related thereto), and the  
25 acquired information, determining whether the subject has a VR-3 or VR-5-associated disease or disorder or a pre-disposition to a VR-3 or VR-5-associated disease or disorder. The method may further comprise the step of recommending a particular treatment for the disease, disorder or pre-disease condition.

          The present invention also provides a business method for determining whether a  
30 subject has a VR-3 or VR-5-associated disease or disorder or a pre-disposition to a VR-3 or VR-5-associated disease or disorder, said method comprising the steps of receiving information related to VR-3 or VR-5 (*e.g.*, sequence information and/or information



related thereto), receiving phenotypic information associated with the subject, acquiring information from the network related to VR-3 or VR-5 and/or related to a VR-3 or VR-5-associated disease or disorder, and based on one or more of the phenotypic information, the VR-3 or VR-5 information, and the acquired information, determining  
5 whether the subject has a VR-3 or VR-5-associated disease or disorder or a pre-disposition to a VR-3 or VR-5-associated disease or disorder. The method may further comprise the step of recommending a particular treatment for the disease, disorder or pre-disease condition.

The invention also includes an array comprising a VR-3 or VR-5 sequence of the  
10 present invention. The array can be used to assay expression of one or more genes in the array. In one embodiment, the array can be used to assay gene expression in a tissue to ascertain tissue specificity of genes in the array. In this manner, up to about 7600 genes can be simultaneously assayed for expression, one of which can be VR-3 or VR-5. This allows a profile to be developed showing a battery of genes specifically expressed  
15 in one or more tissues.

In addition to such qualitative determination, the invention allows the quantitation of gene expression. Thus, not only tissue specificity, but also the level of expression of a battery of genes in the tissue is ascertainable. Thus, genes can be grouped on the basis of their tissue expression *per se* and level of expression in that  
20 tissue. This is useful, for example, in ascertaining the relationship of gene expression between or among tissues. Thus, one tissue can be perturbed and the effect on gene expression in a second tissue can be determined. In this context, the effect of one cell type on another cell type in response to a biological stimulus can be determined. Such a determination is useful, for example, to know the effect of cell-cell interaction at the  
25 level of gene expression. If an agent is administered therapeutically to treat one cell type but has an undesirable effect on another cell type, the invention provides an assay to determine the molecular basis of the undesirable effect and thus provides the opportunity to co-administer a counteracting agent or otherwise treat the undesired effect. Similarly, even within a single cell type, undesirable biological effects can be  
30 determined at the molecular level. Thus, the effects of an agent on expression of other than the target gene can be ascertained and counteracted.

In another embodiment, the array can be used to monitor the time course of expression of one or more genes in the array. This can occur in various biological contexts, as disclosed herein, for example development of a VR-3 or VR-5-associated disease or disorder, progression of VR-3 or VR-5-associated disease or disorder, and  
5 processes, such a cellular transformation associated with the VR-3 or VR-5-associated disease or disorder.

The array is also useful for ascertaining the effect of the expression of a gene on the expression of other genes in the same cell or in different cells (*e.g.*, ascertaining the effect of VR-3 or VR-5 expression on the expression of other genes). This provides, for  
10 example, for a selection of alternate molecular targets for therapeutic intervention if the ultimate or downstream target cannot be regulated.

The array is also useful for ascertaining differential expression patterns of one or more genes in normal and abnormal cells. This provides a battery of genes (*e.g.*, including VR-3 or VR-5) that could serve as a molecular target for diagnosis or  
15 therapeutic intervention.

This invention is further illustrated by the following examples which should not be construed as limiting. The contents of all references, patents and published patent applications cited throughout this application, as well as the Figures and the Sequence  
Listing, are incorporated herein by reference.

20

## EXAMPLES

### **EXAMPLE 1: IDENTIFICATION AND CHARACTERIZATION OF HUMAN VR-3 AND VR-5 CDNA**

25 In this example, the identification and characterization of the genes encoding human VR-3 (clone 18615) and human VR-5 (clone 48003) is described.

#### Isolation of the human VR-3 and human VR-5 cDNA

The invention is based, at least in part, on the discovery of human genes  
30 encoding novel proteins, referred to herein as VR-3 and VR-5.

The entire sequence of the human clone 18615 was determined and found to contain an open reading frame termed human "VR-3." The nucleotide sequence encoding the human VR-3 protein is shown in Figure 1 and is set forth as SEQ ID NO:1. The protein encoded by this nucleic acid comprises about 725 amino acids and has the amino acid sequence shown in Figure 1 and set forth as SEQ ID NO:2. The coding region (open reading frame) of SEQ ID NO:1 is set forth as SEQ ID NO:3.

The entire sequence of the human clone 48003 was determined and found to contain an open reading frame termed human "VR-5." The nucleotide sequence encoding the human VR-5 protein is shown in Figure 2 and is set forth as SEQ ID NO:4. The protein encoded by this nucleic acid comprises about 871 amino acids and has the amino acid sequence shown in Figure 2 and set forth as SEQ ID NO:5. The coding region (open reading frame) of SEQ ID NO:4 is set forth as SEQ ID NO:6. Clone 48003, comprising the coding region of human VR-5, was deposited with the American Type Culture Collection (ATCC®), 10801 University Boulevard, Manassas, VA 20110-2209, on June 8, 2000, and assigned Accession No. PTA-2013.

#### Analysis of the Human VR-3 Molecule

A BLASTN 2.0 search against the NR database, using a score of 100 and a word length of 12 (Altschul *et al.* (1990) *J. Mol. Biol.* 215:403) of the nucleotide sequence of human VR-3 revealed that human VR-3 is 85% identical to *Rattus norvegicus* calcium transporter CaT1 (GenBank Accession Number AF160798) over nucleotides 66 to 2236, 80% identical over nucleotides 2242 to 2469, 63% identical over nucleotides 2555 to 2984, and 71% over nucleotides 2797 to 2984. This search further revealed that human VR-3 is 80% identical to the mRNA for rabbit (*Oryctolagus cuniculus*) epithelial calcium channel (ECaC; GenBank Accession Number AJ133128) over nucleotides 280 to 2263. Human VR-3 is also 60% identical to *Rattus norvegicus* vanilloid receptor-like protein 1 (VRL-1) (GenBank Accession Number AF129113) over nucleotides 1427 to 1859. This search further revealed that human VR-3 is 57% identical to the mRNA for *Rattus norvegicus* stretch inducible nonselective channel (SIC) (GenBank Accession Number AB015231) over nucleotides 1412 to 1797. This search further identified a region of the human VR-3 that is 60% identical to the mRNA for *Rattus norvegicus* stretch activated channel 2B (rSAC2b; GenBank Accession

Number AB029330) over nucleotides 1427 to 1859. In addition, this search further indicated that the human VR-3 displays 57% identity to the *Rattus norvegicus* mRNA for vanilloid receptor subtype 1 (VR-1) (GenBank Accession Number AF029310) over nucleotides 1412 to 1797.

5           A BLASTN 2.0 search against the dbEST database, using a score of 100 and a word length of 12 (Altschul *et al.* (1990) *J. Mol. Biol.* 215:403) of the nucleotide sequence of human VR-3 revealed that human VR-3 is 98% identical to nc74a09.s1 NCI\_CGAP\_Pr2 *H. sapiens* cDNA clone IMAGE:783064 (Accession Number AA469437) over nucleotides 2186-2666.

10           A BLASTX 2.0 search against the NRP/protot database, using a wordlength of 3, a score of 100, and a BLOSUM62 matrix, of the translated nucleotide sequence of human VR-3 revealed that VR-3 is 89% identical to the amino acid sequence of *Rattus norvegicus* calcium transporter CaT1 (GenBank Accession Number AF160798) over translated nucleotides 278-2452. This search further indicated that the human VR-3 is  
15 74% identical to the amino acid sequence of *O. cuniculus* epithelial calcium channel (GenBank Accession Number AJ133128) over translated nucleotides 278-2434. An identical BLASTX 2.0 search against the PATENT\_2/gsprot database revealed a 29% amino acid sequence identity between human VR-3 and chicken capsaicin receptor subtype VR1 (Accession Number Y06561) over translated nucleotides 398 to 2155.

20           A search was also performed against the ProDom database resulting in the identification of a portion of the deduced amino acid sequence of human VR-3 (SEQ ID NO:2) which has a 36% identity within ProDom Accession Number PD101189 ("rat vanilloid receptor subtype 1") over residues 100 to 220. Human VR-3 is also 30% identical to ProDom Accession Number PD011151 ("protein olfactory channel/vanilloid  
25 receptor subtype F28H7.10") over residues 257 to 351, and is 32% identical to the same molecule over residues 51 to 129. Human VR-3 further exhibits a 28% identity with ProDom Accession Number PD003230 ("protein channel calcium receptor ionic transmembrane ion transport entry transient") over residues 531 to 610 and a 22% identity over residues 382 to 522. The results of this search are shown in Figure 10.

30           A search for domain consensus sequences was also performed using a database of HMMs (the Pfam database, release 2.1) using the default parameters. The search revealed three ankyrin repeat domains (Pfam Accession Number PF00023) within SEQ

ID NO:2 at residues 78-108, 116-148, and 162-194. The results of this search are shown in Figure 8.

A search was also performed against the Prosite database, and resulted in the identification of an N-glycosylation site at residues 208-211 and at residues 358-361.

- 5 The VR-3 protein was aligned with the rat calcium transporter (Accession Number AF160798) and the rabbit epithelial calcium channel (Accession number AJ133128) using the CLUSTALW (1.74) multiple sequence alignment program. The results of the alignment are set forth in Figure 3.

10 Analysis of the Human VR-5 Molecule

- A BLASTN 2.0 search against the NRN database, using a score of 100 and a word length of 12 (Altschul *et al.* (1990) *J. Mol. Biol.* 215:403) of the nucleotide sequence of human VR-5 revealed that human VR-5 is 70% identical to the mRNA for *Rattus norvegicus* stretch inducible nonselective channel (SIC) (GenBank Accession Number AB015231) over nucleotides 936-2854, 63% identical over nucleotides 2787-3088, and 65% identical over nucleotides 3070-3243. This search further revealed that human VR-5 is 64% identical to *Rattus norvegicus* mRNA for vanilloid receptor subtype 1 (VR-1) (GenBank Accession Number AF029310) over nucleotides 362-2542, and 62% identical to *Rattus norvegicus* vanilloid receptor-like protein 1 (VRL-1) (GenBank Accession Number AF129113) over nucleotides 888-2183, and 65% identical over nucleotides 2073-2479. In addition, the search revealed that the human VR-5 is 62% identical to *Rattus norvegicus* vanilloid receptor-like protein 1 (VRL-1) (GenBank Accession Number AF129112) over nucleotides 465-1347, 61% identical over nucleotides 1130-1987, and 67% identical over nucleotides 2090-2480.

- 25 A BLASTN 2.0 search against the dbEST database, using a score of 100 and a word length of 12 (Altschul *et al.* (1990) *J. Mol. Biol.* 215:403) of the nucleotide sequence of human VR-5 revealed that human VR-5 is 89% identical to mq35a11.y1 Barstead MPLRB1 *Mus musculus* cDNA clone IMAGE:580700 5', (similar to TR:035433 vanilloid receptor subtype 1) (Accession Number AI510567) over  
30 nucleotides 1074-1587.

- 112 -

A BLASTX 2.0 search against the NRP/protot database, using a wordlength of 3, a score of 100, and a BLOSUM62 matrix, of the translated nucleotide sequence of human VR-5 revealed that VR-5 is 51% identical to the amino acid sequence of *Rattus norvegicus* vanilloid receptor subtype 1 (VR-1) (GenBank Accession Number AAC53398) over translated nucleotides 474-2564. This search further indicated that the human VR-5 is 60% identical to the amino acid sequence of *Rattus norvegicus* stretch-inhibitable nonselective channel (SIC) (GenBank Accession Number BAA34942) over translated nucleotides 1116-2696. An identical BLASTX 2.0 search against the PATENT\_2/gsprot database revealed a 50% amino acid sequence identity between human VR-5 and chicken capsaicin receptor subtype VR1 (Accession Number Y06561) over translated nucleotides 522 to 2669. In addition, the human VR-5 demonstrated a 51% identity to rat capsaicin receptor subtype VR1 (Accession Number Y06555) over translated nucleotides 474-2564.

A search was also performed against the ProDom database resulting in the identification of a portion of the deduced amino acid sequence of human VR-5 (SEQ ID NO:5) which has a 56% identity within ProDom Accession Number PD101189 ("rat vanilloid receptor subtype 1") over residues 147 to 365, and a 53% identity within ProDom Accession Number PD137334 ("rat vanilloid receptor subtype 1") over residues 752 to 827. Human VR-5 is also 34% identical to ProDom Accession Number PD011151 ("protein olfactory channel/vanilloid receptor subtype F28H7.10") over residues 367 to 498. Human VR-5 further exhibits a 24% identity with ProDom Accession Number PD003230 ("protein channel calcium receptor ionic transmembrane ion transport entry transient") over residues 561 to 635. The results of this search are shown in Figure 11.

A search for domain consensus sequences was also performed using a database of HMMs (the Pfam database, release 2.1) using the default parameters. The search revealed three ankyrin repeat domains (Pfam Accession Number PF00023) within SEQ ID NO:5 at residues 237-269, 284-319, and 369-400. This search also revealed an ion transport protein (Pfam Accession Number PF00520) within SEQ ID NO:5 at residues 473-718. The results of this search are shown in Figure 9.

A search was also performed against the Prosite database, and resulted in the identification of five N-glycosylation sites (at residues 201 to 204, 207 to 210, 651 to 654, 784 to 787, and 802 to 805).

The VR-5 protein was aligned with the amino acid sequence of the *Mus musculus* ion channel (Accession Number AB021875) using the GAP program in the GCG software package (Blosum 62 matrix) and a gap weight of 12 and a length weight of 4. The results of the alignment are set forth in Figure 4. The VR-5 nucleotide sequence was also aligned with that of the *Mus musculus* ion channel (Accession Number AB021875) using the CLUSTAL (1.74) multiple sequence alignment. The results of this alignment are set forth in Figure 5.

#### Tissue Distribution of human VR-3 mRNA by PCR analysis

The tissue distribution of VR-3 mRNA was determined by Polymerase Chain Reaction (PCR) on cDNA libraries using oligonucleotide primers based on the human VR-3 sequence.

The human VR-3 gene was found to be predominantly expressed in placenta, mammary gland, esophagus, a Burkitt's Lymphoma cell line, fetal liver, an acute promyelocyte leukemia (HL60) cell, fetal kidney, thyroid, prostate, and salivary gland. In addition, some expression was also noted in HEVECL cells, bronchial epithelium, erythroleukemia cells, trachea, testes, T24 CTL cells, prostate epithelium, MCP-1 Mast cell line, a lymphoma B cell line (ST486), umbilical smooth muscle, T cells, fetal adrenal gland, fetal lung, mid-term placenta, pulmonary artery smooth muscle, fetal brain, and skin/adipose tissue.

#### Tissue Distribution of human VR-3 mRNA using Taqman™ analysis

This example describes the tissue distribution of human VR-3 mRNA in a variety of cells and tissues, as determined using the TaqMan™ procedure. The Taqman™ procedure is a quantitative, reverse transcription PCR-based approach for detecting mRNA. The RT-PCR reaction exploits the 5' nuclease activity of AmpliTaq Gold™ DNA Polymerase to cleave a TaqMan™ probe during PCR. Briefly, cDNA was generated from the samples of interest, e.g., various human, monkey and rat tissue samples, and used as the starting material for PCR amplification. In addition to the 5'

and 3' gene-specific primers, a gene-specific oligonucleotide probe (complementary to the region being amplified) was included in the reaction (*i.e.*, the Taqman™ probe). The TaqMan™ probe includes the oligonucleotide with a fluorescent reporter dye covalently linked to the 5' end of the probe (such as FAM (6-carboxyfluorescein), TET (6-carboxy-4,7,2',7'-tetrachlorofluorescein), JOE (6-carboxy-4,5-dichloro-2,7-dimethoxyfluorescein), or VIC) and a quencher dye (TAMRA (6-carboxy-N,N,N',N'-tetramethylrhodamine) at the 3' end of the probe.

During the PCR reaction, cleavage of the probe separates the reporter dye and the quencher dye, resulting in increased fluorescence of the reporter. Accumulation of PCR products is detected directly by monitoring the increase in fluorescence of the reporter dye. When the probe is intact, the proximity of the reporter dye to the quencher dye results in suppression of the reporter fluorescence. During PCR, if the target of interest is present, the probe specifically anneals between the forward and reverse primer sites. The 5'-3' nucleolytic activity of the AmpliTaq™ Gold DNA Polymerase cleaves the probe between the reporter and the quencher only if the probe hybridizes to the target. The probe fragments are then displaced from the target, and polymerization of the strand continues. The 3' end of the probe is blocked to prevent extension of the probe during PCR. This process occurs in every cycle and does not interfere with the exponential accumulation of product. RNA was prepared using the trizol method and treated with DNase to remove contaminating genomic DNA. cDNA was synthesized using standard techniques. Mock cDNA synthesis in the absence of reverse transcriptase resulted in samples with no detectable PCR amplification of the control gene confirms efficient removal of genomic DNA contamination.

A normal human phase I panel indicated highest expression of VR-3 mRNA in placenta, followed by salivary gland and prostate tissue. Weak expression was detected in brain, testes, spinal cord, and kidney with weakest expression indicated in heart, mammary gland, small intestine, thymus, trachea, and skin.

A second human phase I panel which included both normal human tissues and human tumor tissues indicated highest expression in cortex, followed by prostate tumor and normal prostate, with higher expression in prostate tumor as compared to normal prostate. Weaker expression was detected in skin followed by hypothalamus, thymus,



- 115 -

spinal cord, kidney, liver, breast tumor and normal breast, with higher expression in breast tumor than in normal breast tissue.

A phase 1.3.2 human tissue panel indicated highest expression in pancreas followed by prostate tumor and normal prostate with greater expression in prostate tumor as compared to normal prostate. Weak expression was detected in breast tumor followed by kidney, liver fibrosis, hypothalamus, cortex, spinal cord, and normal skin. Weaker expression was detected in normal breast, colon tumor, normal liver, and decubitus skin.

A phase 1.5.2 human tissue panel indicated highest expression in pancreas, followed by cortex and prostate tumor. Weaker expression was indicated in normal prostate, salivary glands, and kidney with comparatively less expression detected in hypothalamus, breast tumor, liver fibrosis, and erythroid tissue. Weaker expression was noted in normal skin, spinal cord, normal breast, brain tumor, and lung tumor. Still weaker expression was detected in normal artery, diseased aorta, hemangioma, ovary, normal colon, colon tumor, normal lung, chronic obstructive pulmonary disease (COPD) lung, inflammatory colon disease (IBD) colon, normal liver, tonsil, lymph node, decubitus skin, activated peripheral blood mononuclear cells (PBMC), and megakaryocytes.

Normal tissues tested also included an array of monkey and human tissues. Expression was greatest in monkey cortex and human brain, followed by monkey hairy skin, human and monkey kidney, and human spinal cord. Weak expression was noted in monkey spinal cord, human heart, and human and monkey liver.

A human cardiovascular organ panel was also tested indicating expression in kidney and kidney HT, with weaker expression in Wilms tumor, spinal cord, and liver. Weakest expression was detected in normal atrium.

A CNS rat phase I panel indicated highest expression of VR-3 mRNA in spinal cord followed by hairy skin, brain, striatum, cortex, superior cervical ganglia (SCG), sciatic nerve, brain stem, ipsilateral trigeminal (TRG), thalamus, DRG, dorsal nuclei, and cerebellum. Weaker expression was also detected in lung, followed by heart and liver.

Tissue Distribution of human VR-5 mRNA by PCR analysis

The tissue distribution of VR-5 mRNA was determined by Polymerase Chain Reaction (PCR) on cDNA libraries using oligonucleotide primers based on the human VR-5 sequence.

5        The human VR-5 gene was found to be predominantly expressed in HEVECL cells, trachea cells, mammary gland cells, embryonic keratinocytes, astrocytes, fetal spleen cells, SCC25 CDDP cells (derived from a tongue squamous cell carcinoma), prostate epithelium cells, esophagus, bone marrow, keratinocytes, fetal kidney, thyroid, fetal skin, pulmonary artery smooth muscle, kidney, CHT127 (a colon to liver  
10 metastasis), HUVEC TGF- $\beta$ , HUVEC control cells (umbilical epithelia), Hep-G2 (insulinoma), skin, normal breast epithelia, spleen, normal ovarian epithelia, MCF-7 H (a mammary carcinoma), prostate tumor, lung squamous cell carcinoma (PIT299), HUVEC hypoxia (umbilical epithelial), d8 dendritic, salivary gland, and melanoma (G361 cell line). In addition, some expression was also noted in Burkitt's Lymphoma  
15 cells, A2780 ADR, ME 180 control, pituitary cells, prostate fibroblast, uterine smooth muscle, MCP-1 (a mast cell line), ST486 (lymphoma B cell), primary osteoblast, prostate smooth muscle, fetal liver, HL-60 (acute promyelocytic leukemia), skeletal muscle, osteoblasts, prostate cancer liver metastasis (JHH4), colon to liver metastasis (CHT221), mammary gland, bone marrow (CD34+), ovarian ascites, lung squamous cell  
20 carcinoma (MDA 261), normal prostate, megakaryocytes, and Hepatitis B virus-expressing Hep-G2 cells.

Tissue Distribution of human VR-5 mRNA using Taqman™ analysis

To further investigate the expression of VR-5 in various tissues, TaqMan  
25 analysis was utilized, as described above.

A normal human tissue panel was tested indicating highest expression of VR-5 in kidney. Lesser expression was also noted across a broad array of tissues, including placenta, followed by testes, differentiated and undifferentiated osteoblasts, liver, fetal liver, osteoclasts, tonsil, prostate, spleen, lung, heart, and thyroid.

30        A phase 1.1.3 human tissue panel was also tested. Highest expression was detected in kidney. Weaker expression was detected in a broad array of tissues including HUVEC, followed by aortic SMCs, late aortic SMCs, liver fibrosis, HMVEC,

fetal liver, epithelial cells, early aortic SMCs, and epithelial cells, undifferentiated osteoblasts, differentiated osteoblasts, primary osteoblasts, osteoclasts, endothelial cells, ovary tumor, prostate tumor, normal breast, breast tumor, liver, pancreas normal, prostate tumor, heart, and fetal heart.

5           A phase 1 panel indicated highest expression in kidney, followed by prostate epithelial cells, endothelial cells, and primary osteoblasts. Weaker expression was detected in a broad array of tissues, including liver fibrosis and normal fetal liver followed by ovary tumor, normal ovary (with ovary tumor expression higher than normal ovary), normal prostate, prostate tumor, normal breast, breast tumor, cortex, 10   hypothalamus, nerve, and differentiated and undifferentiated osteoblasts.

A cardiovascular organ panel indicated highest expression in normal human kidney, followed by kidney HT NDR and kidney HT CHT. Expression was also detected in monkey heart and human skeletal muscle.

A second cardiovascular organ panel indicated highest expression in kidney, 15   followed by kidney HT, then liver.

Expression analysis of clinical lung samples indicated expression of VR-5 mRNA in lung tumor samples (*e.g.*, CHT 911, MDA 262, CHT 814, and CHT 726 lung tumor samples), but negative expression in normal lung samples.

Higher expression was also detected ovarian tumor samples tested, as compared 20   to normal ovary samples.

Furthermore, higher expression of VR-5 was detected in breast tumor samples tested, as compared to normal breast samples.

A second oncology panel was tested which indicated highest expression in Wilms tumor, followed by placenta, normal liver, HMVEC Prol, HMVEC Arr, liver 25   metastasis, fetal liver, fetal liver, renal tumor, and endometrium. Expression in normal liver was higher than expression in liver metastasis.

Expression in MCF10A and MCF3B EGF treated cells was analyzed over eight hours. Expression of VR-5 mRNA increased between 0 and 2 hours, and decreased from 2 hours to 8 hours in MCF3B EGF treated cells. Expression in MCF10A EGF 30   treated cells did not significantly increase or decrease over 8 hours.

Expression in MCF10A variant cells indicated highest expression in MCF10CA1a.cl1-agar, followed by MCF10CA1a.cl1-T, and MCF10MS-NT.

Analysis of a human and monkey tissue panel indicated the highest expression of VR-5 in kidney from monkey and human. Expression of VR-5 was also noted in human and monkey liver, human brain, and monkey hairy skin.

5       **EXAMPLE 2:       EXPRESSION OF RECOMBINANT VR-3 AND VR-5  
                          PROTEIN IN BACTERIAL CELLS**

In this example, VR-3 or VR-5 is expressed as a recombinant glutathione-S-transferase (GST) fusion polypeptide in *E. coli* and the fusion polypeptide is isolated  
10 and characterized. Specifically, VR-3 or VR-5 is fused to GST and this fusion polypeptide is expressed in *E. coli*, e.g., strain PEB199. Expression of the GST-VR-3 or GST-VR-5 fusion protein in PEB199 is induced with IPTG. The recombinant fusion polypeptide is purified from crude bacterial lysates of the induced PEB199 strain by affinity chromatography on glutathione beads. Using polyacrylamide gel  
15 electrophoretic analysis of the polypeptide purified from the bacterial lysates, the molecular weight of the resultant fusion polypeptide is determined.

**EXAMPLE 3:       EXPRESSION OF RECOMBINANT VR-3 AND VR-5  
                          PROTEIN IN COS CELLS**

20

To express the VR-3 or VR-5 gene in COS cells, the pcDNA/Amp vector by Invitrogen Corporation (San Diego, CA) is used. This vector contains an SV40 origin of replication, an ampicillin resistance gene, an *E. coli* replication origin, a CMV promoter followed by a polylinker region, and an SV40 intron and polyadenylation site. A DNA  
25 fragment encoding the entire VR-3 or VR-5 protein and an HA tag (Wilson *et al.* (1984) *Cell* 37:767) or a FLAG tag fused in-frame to its 3' end of the fragment is cloned into the polylinker region of the vector, thereby placing the expression of the recombinant protein under the control of the CMV promoter.

To construct the plasmid, the VR-3 or VR-5 DNA sequence is amplified by PCR  
30 using two primers. The 5' primer contains the restriction site of interest followed by approximately twenty nucleotides of the VR-3 or VR-5 coding sequence starting from the initiation codon; the 3' end sequence contains complementary sequences to the other

restriction site of interest, a translation stop codon, the HA tag or FLAG tag and the last 20 nucleotides of the VR-3 or VR-5 coding sequence. The PCR amplified fragment and the pCDNA/Amp vector are digested with the appropriate restriction enzymes and the vector is dephosphorylated using the CIAP enzyme (New England Biolabs, Beverly, MA). Preferably the two restriction sites chosen are different so that the VR-3 or VR-5 gene is inserted in the correct orientation. The ligation mixture is transformed into *E. coli* cells (strains HB101, DH5 $\alpha$ , SURE, available from Stratagene Cloning Systems, La Jolla, CA, can be used), the transformed culture is plated on ampicillin media plates, and resistant colonies are selected. Plasmid DNA is isolated from transformants and examined by restriction analysis for the presence of the correct fragment.

COS cells are subsequently transfected with the VR-3- or VR5-pcDNA/Amp plasmid DNA using the calcium phosphate or calcium chloride co-precipitation methods, DEAE-dextran-mediated transfection, lipofection, or electroporation. Other suitable methods for transfecting host cells can be found in Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual*. 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989. The expression of the VR-3 or VR-5 polypeptide is detected by radiolabelling (<sup>35</sup>S-methionine or <sup>35</sup>S-cysteine available from NEN, Boston, MA, can be used) and immunoprecipitation (Harlow, E. and Lane, D. *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1988) using an HA specific monoclonal antibody. Briefly, the cells are labelled for 8 hours with <sup>35</sup>S - methionine (or <sup>35</sup>S -cysteine). The culture media are then collected and the cells are lysed using detergents (RIPA buffer, 150 mM NaCl, 1% NP-40, 0.1% SDS, 0.5% DOC, 50 mM Tris, pH 7.5). Both the cell lysate and the culture media are precipitated with an HA specific monoclonal antibody. Precipitated polypeptides are then analyzed by SDS-PAGE.

Alternatively, DNA containing the VR-3 or VR-5 coding sequence is cloned directly into the polylinker of the pCDNA/Amp vector using the appropriate restriction sites. The resulting plasmid is transfected into COS cells in the manner described above, and the expression of the VR-3 or VR-5 polypeptide is detected by radiolabelling and immunoprecipitation using a VR-3 or VR-5 specific monoclonal antibody.

**Equivalents**

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following

5 claims.

**What is claimed:**

1. An isolated nucleic acid molecule selected from the group consisting of:
  - a) a nucleic acid molecule comprising the nucleotide sequence set forth in  
5 SEQ ID NO:1 or SEQ ID NO:4; and
  - b) a nucleic acid molecule comprising the nucleotide sequence set forth  
in SEQ ID NO:3 or SEQ ID NO:6.
2. An isolated nucleic acid molecule which encodes a polypeptide  
10 comprising the amino acid sequence set forth in SEQ ID NO:2 or SEQ ID NO:5.
3. An isolated nucleic acid molecule comprising the nucleotide sequence  
contained in the plasmid deposited with ATCC® as Accession Number PTA-2013.
- 15 4. An isolated nucleic acid molecule which encodes a naturally occurring  
allelic variant of a polypeptide comprising the amino acid sequence set forth in SEQ ID  
NO:2 or SEQ ID NO:5.
5. An isolated nucleic acid molecule selected from the group consisting of:
  - 20 a) a nucleic acid molecule comprising a nucleotide sequence which is at  
least 91% identical to the nucleotide sequence of SEQ ID NO:1, SEQ ID NO:3, SEQ ID  
NO:4, or SEQ ID NO:6 or a complement thereof;
  - b) a nucleic acid molecule comprising a fragment of at least 15  
nucleotides of a nucleic acid comprising the nucleotide sequence of SEQ ID NO:1, SEQ  
25 ID NO:3, SEQ ID NO:4, or SEQ ID NO:6, or a complement thereof;
  - c) a nucleic acid molecule which encodes a polypeptide comprising an  
amino acid sequence at least about 91% identical to the amino acid sequence of SEQ ID  
NO:2 or SEQ ID NO:5; and
  - d) a nucleic acid molecule which encodes a fragment of a polypeptide  
30 comprising the amino acid sequence of SEQ ID NO:2 or SEQ ID NO:5, wherein the  
fragment comprises at least 15 contiguous amino acid residues of the amino acid  
sequence of SEQ ID NO:2 or SEQ ID NO:5.

6. An isolated nucleic acid molecule which hybridizes to a complement of the nucleic acid molecule of any one of claims 1, 2, 3, 4, or 5 under stringent conditions.
7. An isolated nucleic acid molecule comprising a nucleotide sequence  
5 which is complementary to the nucleotide sequence of the nucleic acid molecule of any one of claims 1, 2, 3, 4, or 5.
8. An isolated nucleic acid molecule comprising the nucleic acid molecule of any one of claims 1, 2, 3, 4, or 5, and a nucleotide sequence encoding a heterologous  
10 polypeptide.
9. A vector comprising the nucleic acid molecule of any one of claims 1, 2, 3, 4, or 5.
10. The vector of claim 9, which is an expression vector.  
15
11. A host cell transfected with the expression vector of claim 10.
12. A method of producing a polypeptide comprising culturing the host cell  
20 of claim 11 in an appropriate culture medium to, thereby, produce the polypeptide.
13. An isolated polypeptide selected from the group consisting of:  
a) a fragment of a polypeptide comprising the amino acid sequence of SEQ ID NO:2 or SEQ ID NO:5, wherein the fragment comprises at least 10 contiguous  
25 amino acids of SEQ ID NO:2 or SEQ ID NO:5;  
b) a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NO:2 or SEQ ID NO:5, wherein the polypeptide is encoded by a nucleic acid molecule which hybridizes to a complement of a nucleic acid molecule consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:4, or SEQ ID NO:6  
30 under stringent conditions;



c) a polypeptide which is encoded by a nucleic acid molecule comprising a nucleotide sequence which is at least 91% identical to a nucleic acid comprising the nucleotide sequence of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:4, or SEQ ID NO:6; and

5 d) a polypeptide comprising an amino acid sequence which is at least 91% identical to the amino acid sequence of SEQ ID NO:2 or SEQ ID NO:5.

14. The isolated polypeptide of claim 13 comprising the amino acid sequence of SEQ ID NO:2 or SEQ ID NO:5.

10

15. The polypeptide of claim 13, further comprising heterologous amino acid sequences.

16. An antibody which selectively binds to a polypeptide of claim 13.

15

17. A method for detecting the presence of a polypeptide of claim 13 in a sample comprising:

a) contacting the sample with a compound which selectively binds to the polypeptide; and

20

b) determining whether the compound binds to the polypeptide in the sample to thereby detect the presence of a polypeptide of claim 13 in the sample.

18. The method of claim 17, wherein the compound which binds to the polypeptide is an antibody.

25

19. A kit comprising a compound which selectively binds to a polypeptide of claim 13 and instructions for use.

20. A method for detecting the presence of a nucleic acid molecule of any one of claims 1, 2, 3, 4, or 5 in a sample comprising:

a) contacting the sample with a nucleic acid probe or primer which selectively hybridizes to a complement of the nucleic acid molecule; and

5 b) determining whether the nucleic acid probe or primer binds to the complement of the nucleic acid molecule in the sample to thereby detect the presence of the nucleic acid molecule of any one of claims 1, 2, 3, 4, or 5 in the sample.

21. The method of claim 20, wherein the sample comprises mRNA  
10 molecules and is contacted with a nucleic acid probe.

22. A kit comprising a compound which selectively hybridizes to a complement of the nucleic acid molecule of any one of claims 1, 2, 3, 4, or 5 and instructions for use.

15

23. A method for identifying a compound which binds to a polypeptide of claim 13 comprising:

a) contacting the polypeptide, or a cell expressing the polypeptide with a test compound; and

20 b) determining whether the polypeptide binds to the test compound.

24. The method of claim 23, wherein the binding of the test compound to the polypeptide is detected by a method selected from the group consisting of:

a) detection of binding by direct detection of test compound/polypeptide  
25 binding;

b) detection of binding using a competition binding assay; and

c) detection of binding using an assay for VR-3 or VR-5 activity.

25. A method for modulating the activity of a polypeptide of claim 13  
30 comprising contacting the polypeptide or a cell expressing the polypeptide with a compound which binds to the polypeptide in a sufficient concentration to modulate the activity of the polypeptide.

26. A method for identifying a compound which modulates the activity of a polypeptide of claim 13 comprising:

a) contacting the polypeptide, or a cell expressing the polypeptide with a test compound; and

5 b) determining the effect of the test compound on the activity of the polypeptide to thereby identify a compound which modulates the activity of the polypeptide.

27. The method of claim 26, wherein said activity is pain signaling.

10

28. The method of claim 26, wherein said activity is cellular growth and proliferation.

29. A method for identifying a compound which modulates pain comprising:

15

a) contacting the polypeptide of claim 13, or a cell expressing the polypeptide with a test compound; and

b) identifying the compound as a modulator of pain by determining the effect of the test compound on the activity of the polypeptide.

20 30. A method for identifying a compound which modulates pain signaling comprising:

a) contacting the polypeptide of claim 13, or a cell expressing the polypeptide with a test compound; and

25 b) identifying the compound as a modulator of pain signaling by determining the effect of the test compound on the activity of the polypeptide.

31. A method for treating a subject having a pain disorder comprising administering to the subject a VR-3 or VR-5 modulator, thereby treating said subject having a pain disorder.

30

32. A method for treating a subject having a pain disorder comprising administering to the subject a VR-3 or VR-5 modulator, wherein the VR-3 or VR-5 modulator is a modulator identified by the method of claim 26, thereby treating said subject having a pain disorder.

5

33. The method of claim 31, wherein the VR-3 or VR-5 modulator is selected from the group consisting of a small molecule, an antibody specific for VR-3 or VR-5, a VR-3 or VR-5 polypeptide, a fragment of a VR-3 or VR-5 polypeptide, a VR-3 or VR-5 nucleic acid molecule, a fragment of a VR-3 or VR-5 nucleic acid molecule, an  
10 antisense VR-3 or VR-5 nucleic acid molecule, and a ribozyme.

34. The method of claim 31, wherein said VR-3 or VR-5 modulator is administered in a pharmaceutically acceptable formulation.

15 35. The method of claim 31, wherein said VR-3 or VR-5 modulator is administered using a gene therapy vector.

36. A method for identifying a compound which modulates cellular growth or proliferation comprising:

20 a) contacting the polypeptide of claim 13, or a cell expressing the polypeptide with a test compound; and

b) identifying the compound as a modulator of cellular growth or proliferation by determining the effect of the test compound on the activity of the polypeptide.

25

37. A method for treating a subject having a cellular growth or proliferation disorder comprising administering to the subject a VR-3 or VR-5 modulator, thereby treating said subject having a cellular growth or proliferation disorder.

38. A method for treating a subject having a cellular growth or proliferation disorder comprising administering to the subject a VR-3 or VR-5 modulator, wherein the VR-3 or VR-5 modulator is a modulator identified by the method of claim 36, thereby treating said subject having a cellular growth or proliferation disorder.

5

39. The method of claim 37, wherein the VR-3 or VR-5 modulator is selected from the group consisting of a small molecule, an antibody specific for VR-3 or VR-5, a VR-3 or VR-5 polypeptide, a fragment of a VR-3 or VR-5 polypeptide, a VR-3 or VR-5 nucleic acid molecule, a fragment of a VR-3 or VR-5 nucleic acid molecule, an antisense VR-3 or VR-5 nucleic acid molecule, and a ribozyme.

10

40. The method of claim 37, wherein said VR-3 or VR-5 modulator is administered in a pharmaceutically acceptable formulation.

15

41. The method of claim 37, wherein said VR-3 or VR-5 modulator is administered using a gene therapy vector.

42. A method of inhibiting tumor progression in a subject comprising administering to said subject a VR-3 or VR-5 inhibitor such that tumor progression is inhibited in said subject.

20

43. A method of identifying a subject having a cellular growth or proliferation disorder, or at risk for developing a cellular growth or proliferation disorder comprising:

25

a) contacting a sample obtained from said subject comprising polypeptides with a VR-3 or VR-5 binding substance; and

b) detecting the presence of a polypeptide in said sample that binds to said VR-3 or VR-5 binding substance, thereby identifying a subject having a cellular growth or proliferation disorder, or at risk for developing a cellular growth or proliferation disorder.

30

1/22

GATC NAG N T N T C C N T C T T G T N A A C C T C G T A C G C C N C G C G G A A T T C C C G G G T C G A C C C A C G C G T C C G G C C A A G T G T A A C  
 A A A C T C A C A G C C C T C T C C A A A C T G G C T G G G G C T G C T G G G A G A C T C C C A A G G A A C T C G T C A G G A A G G C A G G A G A C A G G A G  
 A C G G G A C C T C T A C A G G G A G A C G G T G G G C C G G C C T T G G G G G G C T G A T G T G G C C C C A A G G C T G A G T C C C G T C A G G G T C T G  
  
 M G L S L P K E K 9  
 G C C T C G G C C T C A G G C C C C C A A G G A G C C G G C C T A C A C C C C A T G G G T T G T C A C T G C C C A A G G A G A A A 27  
  
 G L I L C L W S K F C R W F Q R R E S W 29  
 G G G C T A A T T C T C T G C C T A T G G A G C A A G T T C T G C A G A T G G T T C C A G A G A C G G G A G T C C T G G 87  
  
 A Q S R D E Q N L L Q Q K R I W E S P L 49  
 G C C C A G A G C C G A G A T G A G C A G A A C C T G C T G C A G C A G A A G A G G A T C T G G G A G T C T C C T C T C 147  
  
 L L A A K D N D V Q A L N K L L K Y E D 69  
 C T T C T A G C T G C C A A A G A T A A T G A T G T C C A G G C C C T G A A C A A G T T G C T C A A G T A T G A G G A T 207  
  
 C K V H Q R G A M G E T A L H I A A L Y 89  
 T C C A A G G T G C A C C A G A G A G G A G C C A T G G G G A A A C A G C G C T A C A C A T A G C A G C C C T C T A T 267  
  
 D N L E A A M V L M E A A P E L V F E P 109  
 G A C A A C C T G G A G G C C G C C A T G G T G C T G A T G G A G G C T G C C C G G A G C T G G T C T T T G A G C C C 327  
  
 M T S E L Y E G Q T A L H I A V V N Q N 129  
 A T G A C A T C T G A G C T C T A T G A G G G T C A G A C T G C A C T G C A C A T C G C T G T T G T G A A C C A G A A C 387  
  
 M N L V R A L L A R R A S V S A R A T G 149  
 A T G A A C C T G G T G C G A G C C C T G C T T G C C C G C G A G G C C A G T G T C T C T C C C A G A G C C A C A G G C 447  
  
 T A F R R S P C N L I Y F G E H P L S F 169  
 A C T G C C T T C C G C C G T A G T C C C T G C A A C C T C A T C T A C T T T G G G G A G C A C C C T T T G T C C T T T 507  
  
 A A C V N S E E I V R L L I E H G A D I 189  
 G C T G C C T G T G T G A A C A G T G A G G A G A T C G T G C G C T G C T C A T T G A G C A T G G A G C T G A C A T C 567  
  
 R A Q D S L G N T V L H I L I L Q P N K 209  
 C G G G C C C A G G A C T C C T G G G A A A C A C A G T G T T A C A C A T C C T C A T C C T C C A G C C C A A C A A A 627  
  
 T F A C Q M Y N L L L S Y D R H G D H L 229  
 A C C T T T G C C T G C C A G A T G T A C A A C C T G T T G C T G T C C T A C G A C A G A C A T G G G G A C C A C C T G 687  
  
 Q P L D L V P N H Q G L T P F K L A G V 249  
 C A G C C C C T G G A C C T C G T G C C C A A T C A C C A G G G T C T C A C C C C T T T C A A G C T G G C T G G A G T G 747  
  
 E G N T V M F Q H L M Q K R K H T Q W T 269  
 G A G G G T A A C A C T G T G A T G T T T C A G C A C C T G A T G C A G A A G C G G A A G C A C A C C C A G T G G A C G 807  
  
 Y G P L T S T L Y D L T E I D S S G D E 289  
 T A T G G A C C A C T G A C C T C G A C T C T C T A T G A C C T C A C A G A T C G A C T C C T C A G G G G A T G A G 867  
  
 Q S L L E L I I T T K K R E A R Q I L D 309  
 C A G T C C C T G C T G G A A C T T A T C A T C A C C A C C A A G A A G C G G G A G G C T C G C C A G A T C C T G G A C 927  
  
 Q T P V K E L V S L K W K R Y G R P Y F 329  
 C A G A C G C C G G T G A A G G A G C T G G T G A G C T C A A G T G G A A G C G G T A C G G G C G G C C G T A C T T C 987

Fig. 1A

2/22

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |      |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|
| C   | M   | L   | G   | A   | I   | Y   | L   | L   | Y   | I   | I   | C   | F   | T   | M   | C   | G   | I   | Y   | 349  |
| TGC | ATG | CTG | GGT | GCC | ATA | TAT | CTG | CTG | TAC | ATC | ATC | TGC | TTC | ACC | ATG | TGC | TGC | ATC | TAC | 1047 |
| R   | P   | L   | K   | P   | R   | T   | N   | N   | R   | T   | S   | P   | R   | D   | N   | T   | L   | L   | Q   | 369  |
| CGC | CCC | CTC | AAG | CCC | AGG | ACC | AAT | AAC | CGC | ACG | AGC | CCC | CGG | GAC | AAC | ACC | CTC | TTA | CAG | 1107 |
| Q   | K   | L   | L   | Q   | E   | A   | Y   | M   | T   | P   | K   | D   | D   | I   | R   | L   | V   | G   | E   | 389  |
| CAG | AAG | CTA | CTT | CAG | GAA | GCC | TAC | ATG | ACC | CCT | AAG | GAC | GAT | ATC | CGG | CTG | GTC | GGG | GAG | 1167 |
| L   | V   | T   | V   | I   | G   | A   | I   | I   | I   | L   | L   | V   | E   | V   | P   | D   | I   | F   | R   | 409  |
| CTG | GTG | ACT | GTC | ATT | GGG | GCT | ATC | ATC | ATC | CTG | CTG | GTA | GAG | GTT | CCA | GAC | ATC | TTC | AGA | 1227 |
| M   | G   | V   | T   | R   | F   | F   | G   | Q   | T   | I   | L   | G   | G   | P   | F   | H   | V   | L   | I   | 429  |
| ATG | GGG | GTC | ACT | CGC | TTC | TTT | GGA | CAG | ACC | ATC | CTT | GGG | GGC | CCA | TTC | CAT | GTC | CTC | ATC | 1287 |
| I   | T   | Y   | A   | F   | M   | V   | L   | V   | T   | M   | V   | M   | R   | L   | I   | S   | A   | S   | G   | 449  |
| ATC | ACC | TAT | GCC | TTC | ATG | GTG | CTG | GTG | ACC | ATG | GTG | ATG | CGG | CTC | ATC | AGT | GCC | AGC | GGG | 1347 |
| E   | V   | V   | P   | M   | S   | F   | A   | L   | V   | L   | G   | W   | C   | N   | V   | M   | Y   | F   | A   | 469  |
| GAG | GTG | GTA | CCC | ATG | TCC | TTT | GCA | CTC | GTG | CTG | GGC | TGG | TGC | AAC | GTC | ATG | TAC | TTT | GCC | 1407 |
| R   | G   | F   | Q   | M   | L   | G   | P   | F   | T   | I   | M   | N   | Q   | K   | M   | I   | F   | G   | D   | 489  |
| CGA | GGA | TTC | CAG | ATG | CTA | GGC | CCC | TTC | ACC | ATC | ATG | AAT | CAG | AAG | ATG | ATT | TTT | GGC | GAC | 1467 |
| L   | M   | R   | F   | C   | W   | L   | M   | A   | V   | V   | I   | L   | G   | F   | A   | S   | A   | F   | Y   | 509  |
| CTG | ATG | CGA | TTC | TGC | TGG | CTG | ATG | GCT | GTG | GTC | ATC | CTG | GGC | TTT | GCT | TCA | GCC | TTC | TAT | 1527 |
| I   | I   | F   | Q   | T   | E   | D   | P   | E   | E   | L   | G   | H   | F   | Y   | D   | Y   | P   | M   | A   | 529  |
| ATC | ATC | TTC | CAG | ACA | GAG | GAC | CCC | GAG | GAG | CTA | GGC | CAC | TTC | TAC | GAC | TAC | CCC | ATG | GCC | 1587 |
| L   | F   | S   | T   | F   | E   | L   | F   | L   | T   | I   | I   | D   | G   | P   | A   | N   | Y   | N   | V   | 549  |
| CTG | TTC | AGC | ACC | TTC | GAG | CTG | TTC | CTT | ACC | ATC | ATC | GAT | GGC | CCA | GCC | AAC | TAC | AAC | GTG | 1647 |
| D   | L   | P   | F   | M   | Y   | S   | I   | T   | Y   | A   | A   | F   | A   | I   | I   | A   | T   | L   | L   | 569  |
| GAC | CTG | CCC | TTC | ATG | TAC | AGC | ATC | ACC | TAT | GCT | GCC | TTT | GCC | ATC | ATC | GCC | ACA | CTG | CTC | 1707 |
| M   | L   | N   | L   | L   | I   | A   | M   | M   | G   | D   | T   | H   | W   | R   | V   | A   | H   | E   | R   | 589  |
| ATG | CTC | AAC | CTC | CTC | ATT | GCC | ATG | ATG | GGC | GAC | ACT | CAC | TGG | CGA | GTG | GCC | CAT | GAG | CGG | 1767 |
| D   | E   | L   | W   | R   | A   | Q   | I   | V   | A   | T   | T   | V   | M   | L   | E   | R   | K   | L   | P   | 609  |
| GAT | GAG | CTG | TGG | AGG | GCC | CAG | ATT | GTG | GCC | ACC | ACG | GTG | ATG | CTG | GAG | CGG | AAG | CTG | CCT | 1827 |
| R   | C   | L   | W   | P   | R   | S   | G   | I   | C   | G   | R   | E   | Y   | G   | L   | G   | D   | R   | W   | 629  |
| CGC | TGC | CTG | TGG | CCT | CGC | TCC | GGG | ATC | TGC | GGA | CGG | GAG | TAT | GGC | CTG | GGA | GAC | CGC | TGG | 1887 |
| F   | L   | R   | V   | E   | D   | R   | Q   | D   | L   | N   | R   | Q   | R   | I   | Q   | R   | Y   | A   | Q   | 649  |
| TTC | CTG | CGG | GTG | GAA | GAC | AGG | CAA | GAT | CTC | AAC | CGG | CAG | CGG | ATC | CAA | CGC | TAC | GCA | CAG | 1947 |
| A   | F   | H   | T   | R   | G   | S   | E   | D   | L   | D   | K   | D   | S   | V   | E   | K   | L   | E   | L   | 669  |
| GCC | TTC | CAC | ACC | CGG | GGC | TCT | GAG | GAT | TTG | GAC | AAA | GAC | TCA | GTG | GAA | AAA | CTA | GAG | CTG | 2007 |
| G   | C   | P   | F   | S   | P   | H   | L   | S   | L   | P   | M   | P   | S   | V   | S   | R   | S   | T   | S   | 689  |
| GGC | TGT | CCC | TTC | AGC | CCC | CAC | CTG | TCC | CTT | CCT | ATG | CCC | TCA | GTG | TCT | CGA | AGT | ACC | TCC | 2067 |
| R   | S   | S   | A   | N   | W   | E   | R   | L   | R   | Q   | G   | T   | L   | R   | R   | D   | L   | R   | G   | 709  |
| CGC | AGC | AGT | GCC | AAT | TGG | GAA | AGG | CTT | CGG | CAA | GGG | ACC | CTG | AGG | AGA | GAC | CTG | CGT | GGG | 2127 |
| I   | I   | N   | R   | G   | L   | E   | D   | G   | E   | S   | W   | E   | Y   | Q   | I   | *   |     |     |     | 726  |
| ATA | ATC | AAC | AGG | GGT | CTG | GAG | GAC | GGG | GAG | AGC | TGG | GAA | TAT | CAG | ATC | TGA |     |     |     | 2178 |

Fig. 1B

3/22

CTGCGTGTTCTCACTTCGCTTCCTGGAAC TTGCTCTCATTTTCTGGGTGCATCAAACAAAACAAAACCAAACACCCA  
GAGGTCTCATCTCCCAGGCCCCAGGGGAGAAAGAGGAGTAGCATGAACGCCAAGGAATGTACGTTGAGAATCACTGCTC  
CAGGCCTGCATTACTCCTTCAGCTCTGGGGCAGAGGAAGCCCAGCCCAAGCACGGGGCTGGCAGGGCGTGAGGAACTCT  
CCTGTGGCCTGCTCATCACCTTCCGACAGGAGCACTGCATGTCAGAGCACTTTAAAAACAGGCCAGCCTGCTTGGGCG  
CTCGGTCTCCACCCCAGGGTCATAAGTGGGGAGAGAGCCCTTCCCAGGGCACCCAGGCAGGTGCAGGGGAAGTGCAGAGC  
TTGTGGAAAGCGTGTGAGTGAGGGAGACAGGAACGGCTCTGGGGGTGGGAAGTGGGGCTAGGTCTTGCCAACTCCATCT  
TCAATAAAGTCGTTTTTCGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGAKSCNAGCTTGGACGTTGCCAA  
TCGTCNTGGTGNGTTTTTC

Fig. 1C



4/22

TCGACCCACGCGTCCGGGATCTCCCGGCCGCCGGCGCCAGCCGTCCCGGAGGCTGAGCAGTGCAGACGGGCCTGGGGC

|   |      |
|---|------|
| M A D S S E G P R A G P G E V A E L   | 18   |
| AGGC ATG GCG GAT TCC AGC GAA GGC CCC CGC GCG GGG CCC GGG GAG GTG GCT GAG CTC    | 54   |
| P G D E S G T P G G E A F P L S S L A N   | 38   |
| CCC GGG GAT GAG AGT GGC ACC CCA GGT GGG GAG GCT TTT CCT CTC TCC TCC CTG GCC AAT | 114  |
| L F E G E D G S L S P S P A D A S R P A   | 58   |
| CTG TTT GAG GGG GAG GAT GGC TCC CTT TCG CCC TCA CCG GCT GAT GCC AGT CGC CCT GCT | 174  |
| G P G D G R P N L R M K F Q G A F R K G   | 78   |
| GGC CCA GGC GAT GGG CGA CCA AAT CTG CGC ATG AAG TTC CAG GGC GCC TTC CGC AAG GGG | 234  |
| V P N P I D L L E S T L Y E S S V V P G   | 98   |
| GTG CCC AAC CCC ATC GAT CTG CTG GAG TCC ACC CTA TAT GAG TCC TCG GTG GTG CCT GGG | 294  |
| P K K A P M D S L F D Y G T Y R H H S S   | 118  |
| CCC AAG AAA GCA CCC ATG GAC TCA CTG TTT GAC TAC GGC ACC TAT CGT CAC CAC TCC AGT | 354  |
| D N K R W R K K I I E K Q P Q S P K A P   | 138  |
| GAC AAC AAG AGG TGG AGG AAG AAG ATC ATA GAG AAG CAG CCG CAG AGC CCC AAA GCC CCT | 414  |
| A P Q P P P I L K V F N R P I L F D I V   | 158  |
| GCC CCT CAG CCG CCC CCC ATC CTC AAA GTC TTC AAC CGG CCT ATC CTC TTT GAC ATC GTG | 474  |
| S R G S T A D L D G L L P F L L T H K K   | 178  |
| TCC CGG GGC TCC ACT GCT GAC CTG GAC GGG CTG CTC CCA TTC TTG CTG ACC CAC AAG AAA | 534  |
| R L T D E E F R E P S T G K T C L P K A   | 198  |
| CGC CTA ACT GAT GAG GAG TTT CGA GAG CCA TCT ACG GGG AAG ACC TGC CTG CCC AAG GCC | 594  |
| L L N L S N G R N D T I P V L L D I A E   | 218  |
| TTG CTG AAC CTG AGC AAT GGC CGC AAC GAC ACC ATC CCT GTG CTG CTG GAC ATC GCG GAG | 654  |
| R T G N M R E F I N S P F R D I Y Y R G   | 238  |
| CGC ACC GGC AAC ATG AGG GAG TTC ATT AAC TCG CCC TTC CGT GAC ATC TAC TAT CGA GGT | 714  |
| Q T A L H I A I E R R C K H Y V E L L V   | 258  |
| CAG ACA GCC CTG CAC ATC GCC ATT GAG CGT CGC TGC AAA CAC TAC GTG GAA CTT CTC GTG | 774  |
| A Q G A D V H A Q A R G R F F Q P K D E   | 278  |
| GCC CAG GGA GCT GAT GTC CAC GCC CAG GCC CGT GGG CGC TTC TTC CAG CCC AAG GAT GAG | 834  |
| G G Y F Y F G E L P L S L A A C T N Q P   | 298  |
| GGG GGC TAC TTC TAC TTT GGG GAG CTG CCC CTG TCG CTG GCT GCC TGC ACC AAC CAG CCC | 894  |
| H I V N Y L T E N P H K K A D M R R Q D   | 318  |
| CAC ATT GTC AAC TAC CTG ACG GAG AAC CCC CAC AAG AAG GCG GAC ATG CGG CGC CAG GAC | 954  |
| S R G N T V L H A L V A I A D N T R E N   | 338  |
| TCG CGA GGC AAC ACA GTG CTG CAT GCG CTG GTG GCC ATT GCT GAC AAC ACC CGT GAG AAC | 1014 |
| T K F V T K M Y D L L L L K C A R L F P   | 358  |
| ACC AAG TTT GTT ACC AAG ATG TAC GAC CTG CTG CTG CTC AAG TGT GCC CGC CTC TTC CCC | 1074 |
| D S N L E A V L N N D G L S P L M M A A   | 378  |
| GAC AGC AAC CTG GAG GCC GTG CTC AAC AAC GAC GGC CTC TCG CCC CTC ATG ATG GCT GCC | 1134 |

Fig. 2A

5/22

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |      |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|
| K   | T   | G   | K   | I   | G   | I   | F   | Q   | H   | I   | I   | R   | R   | E   | V   | T   | D   | E   | D   | 398  |
| AAG | ACG | GGC | AAG | ATT | GGG | ATC | TTT | CAG | CAC | ATC | ATC | CGG | CGG | GAG | GTG | ACG | GAT | GAG | GAC | 1194 |
| T   | R   | H   | L   | S   | R   | K   | F   | K   | D   | W   | A   | Y   | G   | P   | V   | Y   | S   | S   | L   | 418  |
| ACA | CGG | CAC | CTG | TCC | CGC | AAG | TTC | AAG | GAC | TGG | GCC | TAT | GGG | CCA | GTG | TAT | TCC | TCG | CTT | 1254 |
| Y   | D   | L   | S   | S   | L   | D   | T   | C   | G   | E   | E   | A   | S   | V   | L   | E   | I   | L   | V   | 438  |
| TAT | GAC | CTC | TCC | TCC | CTG | GAC | ACG | TGT | GGG | GAA | GAG | GCC | TCC | GTG | CTG | GAG | ATC | CTG | GTG | 1314 |
| Y   | N   | S   | K   | I   | E   | N   | R   | H   | E   | M   | L   | A   | V   | E   | P   | I   | N   | E   | L   | 458  |
| TAC | AAC | AGC | AAG | ATT | GAG | AAC | CGC | CAC | GAG | ATG | CTG | GCT | GTG | GAG | CCC | ATC | AAT | GAA | CTG | 1374 |
| L   | R   | D   | K   | W   | R   | K   | F   | G   | A   | V   | S   | F   | Y   | I   | N   | V   | V   | S   | Y   | 478  |
| CTG | CGG | GAC | AAG | TGG | CGC | AAG | TTC | GGG | GCC | GTC | TCC | TTC | TAC | ATC | AAC | GTG | GTC | TCC | TAC | 1434 |
| L   | C   | A   | M   | V   | I   | F   | T   | L   | T   | A   | Y   | Y   | Q   | P   | L   | E   | G   | T   | P   | 498  |
| CTG | TGT | GCC | ATG | GTC | ATC | TTC | ACT | CTC | ACC | GCC | TAC | TAC | CAG | CCG | CTG | GAG | GGC | ACA | CCG | 1494 |
| P   | Y   | P   | Y   | R   | T   | T   | V   | D   | Y   | L   | R   | L   | A   | G   | E   | V   | I   | T   | L   | 518  |
| CCG | TAC | CCT | TAC | CGC | ACC | ACG | GTG | GAC | TAC | CTG | CGG | CTG | GCT | GGC | GAG | GTC | ATT | ACG | CTC | 1554 |
| F   | T   | G   | V   | L   | F   | F   | F   | T   | N   | I   | K   | D   | L   | F   | M   | K   | K   | C   | P   | 538  |
| TTC | ACT | GGG | GTC | CTG | TTC | TTC | TTC | ACC | AAC | ATC | AAA | GAC | TTG | TTC | ATG | AAG | AAA | TGC | CCT | 1614 |
| G   | V   | N   | S   | L   | F   | I   | D   | G   | S   | F   | Q   | L   | L   | Y   | F   | I   | Y   | S   | V   | 558  |
| GGG | GTG | AAT | TCT | CTC | TTC | ATT | GAT | GGC | TCC | TTC | CAG | CTG | CTC | TAC | TTC | ATC | TAC | TCT | GTC | 1674 |
| L   | V   | I   | V   | S   | A   | A   | L   | Y   | L   | A   | G   | I   | E   | A   | Y   | L   | A   | V   | M   | 578  |
| CTG | GTG | ATC | GTC | TCA | GCA | GCC | CTC | TAC | CTG | GCA | GGG | ATC | GAG | GCC | TAC | CTG | GCC | GTG | ATG | 1734 |
| V   | F   | A   | L   | V   | L   | G   | W   | M   | N   | A   | L   | Y   | F   | T   | R   | G   | L   | K   | L   | 598  |
| GTC | TTT | GCC | CTG | GTC | CTG | GGC | TGG | ATG | AAT | GCC | CTT | TAC | TTC | ACC | CGT | GGG | CTG | AAG | CTG | 1794 |
| T   | G   | T   | Y   | S   | I   | M   | I   | Q   | K   | I   | L   | F   | K   | D   | L   | F   | R   | F   | L   | 618  |
| ACG | GGG | ACC | TAT | AGC | ATC | ATG | ATC | CAG | AAG | ATT | CTC | TTC | AAG | GAC | CTT | TTC | CGA | TTC | CTG | 1854 |
| L   | V   | Y   | L   | L   | F   | M   | I   | G   | Y   | A   | S   | A   | L   | V   | S   | L   | L   | N   | P   | 638  |
| CTC | GTC | TAC | TTG | CTC | TTC | ATG | ATC | GGC | TAC | GCT | TCA | GCC | CTG | GTC | TCC | CTC | CTG | AAC | CCG | 1914 |
| C   | A   | N   | M   | K   | V   | C   | N   | E   | D   | Q   | T   | N   | C   | T   | V   | P   | T   | Y   | P   | 658  |
| TGT | GCC | AAC | ATG | AAG | GTG | TGC | AAT | GAG | GAC | CAG | ACC | AAC | TGC | ACA | GTG | CCC | ACT | TAC | CCC | 1974 |
| S   | C   | R   | D   | S   | E   | T   | F   | S   | T   | F   | L   | L   | D   | L   | F   | K   | L   | T   | I   | 678  |
| TCG | TGC | CGT | GAC | AGC | GAG | ACC | TTC | AGC | ACC | TTC | CTC | CTG | GAC | CTG | TTT | AAG | CTG | ACC | ATC | 2034 |
| G   | M   | G   | D   | L   | E   | M   | L   | S   | S   | T   | K   | Y   | P   | V   | V   | F   | I   | I   | L   | 698  |
| GGC | ATG | GGC | GAC | CTG | GAG | ATG | CTG | AGC | AGC | ACC | AAG | TAC | CCC | GTG | GTC | TTC | ATC | ATC | CTG | 2094 |
| L   | V   | T   | Y   | I   | I   | L   | T   | F   | V   | L   | L   | L   | N   | M   | L   | I   | A   | L   | M   | 718  |
| CTG | GTG | ACC | TAC | ATC | ATC | CTC | ACC | TTT | GTG | CTG | CTC | CTC | AAC | ATG | CTC | ATT | GCC | CTC | ATG | 2154 |
| G   | E   | T   | V   | G   | Q   | V   | S   | K   | E   | S   | K   | H   | I   | W   | K   | L   | Q   | W   | A   | 738  |
| GGC | GAG | ACA | GTG | GGC | CAG | GTC | TCC | AAG | GAG | AGC | AAG | CAC | ATC | TGG | AAG | CTG | CAG | TGG | GCC | 2214 |
| T   | T   | I   | L   | D   | I   | E   | R   | S   | F   | P   | V   | F   | L   | R   | K   | A   | F   | R   | S   | 758  |
| ACC | ACC | ATC | CTG | GAC | ATT | GAG | CGC | TCC | TTC | CCC | GTA | TTC | CTG | AGG | AAG | GCC | TTC | CGC | TCT | 2274 |
| G   | E   | M   | V   | T   | V   | G   | K   | S   | S   | D   | G   | T   | P   | D   | R   | R   | W   | C   | F   | 778  |
| GGG | GAG | ATG | GTC | ACC | GTG | GGC | AAG | AGC | TCG | GAC | GGC | ACT | CCT | GAC | CGC | AGG | TGG | TGC | TTC | 2334 |

Fig. 2B

6/22

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |      |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|
| R   | V   | D   | E   | V   | N   | W   | S   |     | W   | N   | Q   | N   | L   | G   | I   | I   | N   | E   | D   | 798  |
| AGG | GTG | GAT | GAG | GTG | AAC | TGG | TCT | CAC | TGG | AAC | CAG | AAC | TTG | GGC | ATC | ATC | AAC | GAG | GAC | 2394 |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |      |
| P   | G   | K   | N   | E   | T   | Y   | Q   | Y   | Y   | G   | F   | S   | H   | T   | V   | G   | R   | L   | R   | 818  |
| CCG | GGC | AAG | AAT | GAG | ACC | TAC | CAG | TAT | TAT | GGC | TTC | TCG | CAT | ACC | GTG | GGC | CGC | CTC | CGC | 2454 |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |      |
| R   | D   | R   | W   | S   | S   | V   | V   | P   | R   | V   | V   | E   | L   | N   | K   | N   | S   | N   | P   | 838  |
| AGG | GAT | CGC | TGG | TCC | TCG | GTG | GTA | CCC | CGC | GTG | GTG | GAA | CTG | AAC | AAG | AAC | TCG | AAC | CCG | 2514 |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |      |
| D   | E   | V   | V   | V   | P   | L   | D   | S   | M   | G   | N   | P   | R   | C   | D   | G   | H   | Q   | Q   | 858  |
| GAC | GAG | GTG | GTG | GTG | CCT | CTG | GAC | AGC | ATG | GGG | AAC | CCC | CGC | TGC | GAT | GGC | CAC | CAG | CAG | 2574 |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |      |
| G   | Y   | P   | R   | K   | W   | R   | T   | D   | D   | A   | P   | L   | *   |     |     |     |     |     |     | 872  |
| GGT | TAC | CCC | CGC | AAG | TGG | AGG | ACT | GAT | GAC | GCC | CCG | CTC | TAG |     |     |     |     |     |     | 2616 |

GGACTGCAGCCCAGCCCCAGCTTCTCTGCCCACTCATTTCTAGTCCAGCCGCAATTCAGCAGTGCCCTTCTGGGGTGTCC  
CCCCACACCTGCTTTGGCCCCAGAGGCGAGGGACCACTGGAGGTGCCAGGGAGGCCCCAGGACCTGTGGTCCCCCTGG  
CTCTGCCTCCCCACCTGGGGTGGGGGCTCCCGGCCACCTGTCTTGCTCCTATGGAGTCACATAAGCCAACGCCAGAGC  
CCCTCCACCTCAGGCCCCAGCCCCTGCCTCTCCATTATTTATTTGCTCTGCTCTCAGGAAGCGACGTGACCCCTGCCCC  
AGCTGGAACCTGGCAGAGGCCTTAGGACCCCGTTCCAAGTGCACTGCCCGGCCAAGCCCCAGCCTCAGCCTGCGCCTGA  
GCTGCATGCGCCACCATTTTTGGCAGCGTGGCAGCTTTGCAAGGGGCTGGGGCCCTCGGCGTGGGGCCATGCCTTCTGT  
GTGTTCTGTAGTGTCTGGGATTTGCCGGTGTCTCAATAAATGGTTATTCAATTGAAAAAAAAAAAAAAAAAAGG

Fig. 2C

7/22

## CLUSTAL W (1.74) multiple sequence alignment

5712756 rat calcium transporter

4581491 rabbit epithelial calcium channel

```

5712756      MGWSLPKEKGLILCLWNKFCRWFRRESWAQSRDEQNLLQOKRIWESPLLLAAKENNVQA
18615        MGLSLPKEKGLILCLWSKFCRWFORRESWAQSRDEQNLLQOKRIWESPLLLAAKDNVQA
4581491      MGACPPKAKGPWAQLQKLLISWPVGEQDWEQYRDRVNMLQOERIRDSPLLLQAAKENDLRL
              ** . ** **      * . : *      .:. * * **. *:***: ** :***** ***:***:

5712756      LIKLLKFEGCEVHQKGAMGETALHTAALYDNLEAAMVLMEAAPELVFEPMTSELYEGQTA
18615        LNKLLKYEDCKVHQKGAMGETALHTAALYDNLEAAMVLMEAAPELVFEPMTSELYEGQTA
4581491      LKILLLNQSCDFQORGAVGETALHVAALYDNLEAATLIMEAAPELAKCEPFGVQTA
              *  **      :.*...*:***:*****:*****:*****:*****. ** . * : ****

5712756      LHTAVINQNVNLVRALLARGASVSARATGSVFHYRPHNLIYYGEHPLSFAACVGSEEIVR
18615        LHTAVVNQNMNLVRALLARRASVSARATGTAFRRSPCNLIYFGEHPLSFAACVNSEEIVR
4581491      LHTAVMNQNLNLVRALLARGASVSARATGAAFRRSPHNLIYYGEHPLSFAACVGSEEIVR
              *****:***:***** *****:.*:  * *****:***** *****

5712756      LLIEHGADIRAQDSLGNLTVLHILILQPNKTFACQMYNLLLSYD-GGDHLKSLELVPNNQG
18615        LLIEHGADIRAQDSLGNLTVLHILILQPNKTFACQMYNLLLSYDRHGDHLQPLDLVPNHQG
4581491      LLIEHGADIRAQDSLGNLTVLHILILQPNKTFACQMYNLLLSYDEHSDHLQSLVLPNHQG
              *****:*****:*****:*****:*****:*****:*****:*****:*****

5712756      LTPFKLAGVEGNIVMFQHLMQKRKHQWITYGPLTSTLYDLTEIDSSGDDQSLLELIVTTK
18615        LTPFKLAGVEGNIVMFQHLMQKRKHTQWITYGPLTSTLYDLTEIDSSGDEQSLLELIITTK
4581491      LTPFKLAGVEGNIVMFQHLMQKRKHVQWTCGPLTSTLYDLTEIDSWGEELSFLVSVSSK
              ***** ***** ***** ***** ***** *: : *:***: : : *

5712756      KREARQILDQTPVKELVSLKWKRYGRPYFCVLGAIYVLYIICFTMCCVYRPLKPRITNRT
18615        KREARQILDQTPVKELVSLKWKRYGRPYFCMLGAIYLLYIICFTMCCYRPLKPRITNRT
4581491      KREARQILEQTPVKELVSFKWKYGRPYFCVLASLYILYMICFTTCCYRPLKLRDDNRT
              *****:*****:***:*****:*. : :*:***: ***** *:***** * ***

5712756      NPRDNILLQOKLLQEAYVTPKDDLRLVGLVSVIGAVIILLVEIPDIFRLGVTRFFGQTI
18615        SPRDNILLQOKLLQEAYMTPKDDIRLVGLVTVIGAIILLVEVPDIFRMGVTRFFGQTI
4581491      DPRDITILLQOKLLQEAYVTHQDNIRLVGLVTVIGAVIILLLEIPDIFRVGASRYFGQTI
              *****:*.*** :*:***** . .*****:*****:*****:*****

5712756      LGGPFHVIIIVTYAFMVLVTMVMRLTNSDGEVVPMSFALVLGWCNVMYFARGFQMLGPFTI
18615        LGGPFHVLIITYAFMVLVTMVMRLISASGEVVPMSFALVLGWCNVMYFARGFQMLGPFTI
4581491      LGGPFHVIIITYASLVLLTMVMRLTNMNGEVVPLSFALVLGWCNVMYFARGFQMLGPFTI
              .*** *:*****: * :*:*****: : *:*****:*****:*. :*:*****

5712756      MIQKMIFGDLMRFCWLMVAVILGFASAFYIIFQTEDPDELGHFYDYPMALESTFELFLTII
18615        MNQKMIFGDLMRFCWLMVAVILGFASAFYIIFQTEDPEELGHFYDYPMALESTFELFLTII
4581491      MIQKMIFGDLMRFCWLMVAVILGFASAFHITFQTEDPNNLGEFSDYPTALFSTFELFLTII
              * *****:*****: * *****:*. * *** *****

5712756      IDGPANYDVDLPFMYSITYAFAI IATLLMLNLLIAMMGDTHWRVAHERDELWRAQVVAT
18615        IDGPANYNDVDLPFMYSITYAFAI IATLLMLNLLIAMMGDTHWRVAHERDELWRAQIVAT
4581491      IDGPANYSDVDLPMYCITYAFAI IATLLMLNLFIAMMGDTHWRVAQERDELWRAQVVAT
              *****:*****:*****:*****:*****:*****:*****:*****

```

Fig. 3A

SUBSTITUTE SHEET (RULE 26)

8/22

5712756 TVMLERKLPRLWPRSGICGREYGLGDRWFLRVEDRQDLNRQIRRYAQAQQDDLYSE  
18615 TVMLERKLPRLWPRSGICGREYGLGDRWFLRVEDRQDLNRQIRRYAQAQFHTRG---SE  
4581491 TVMLERKMPRFLWPRSGICGYEYGLGDRWFLRVENHHDQNPLRVLYVEAFKCS---KE  
\*\*\*\*\*:\*\* \*\*\*\*\* \*\*\*\*\*:::\* \* \*: \*\*.:\*\*:. . .\*

5712756 DLEKDSGEK----LEMAPFGAYLSFPTPSVSRSTSRSSTNWDLRQALRKDLQGIINR  
18615 DLDKDSVEK----LELGCPFSPHLSLPMPSVSRSTSRSSANWERLRQGTLRDLRGIINR  
4581491 DGQEQQLSEKRPSTVESGMLSRSVAFQTPSLSRITTSQSSN--SHRGWEILRRNTLGHLNL  
\* ::: \*\* :\* . . ::: \*\*:\*\*\*:\*\*\* \*\* .: \*\*: \* :\*

5712756 GLEDGEGWEYQI---  
18615 GLEDGESWEYQI---  
4581491 GLDLGEGDGEEVYHF  
\*\*: \*\*. ::

Fig. 3B

9/22

GAP of: fbh48003fl.pep check: 6158 from: 1 to: 871

bh48003FL - Import - complete

to: ab021875.pep check: 1335 from: 1 to: 872

B021875 in GenPept

Symbol comparison table: /ddm\_local/gcg/gcg\_9.1/gcgcore/data/rundata/blosum62  
 CompCheck: 6430

Gap Weight: 12      Average Match: 2.912  
 Length Weight: 4      Average Mismatch: -2.003

Quality: 4128      Length: 872  
 Ratio: 4.739      Gaps: 1  
 Percent Similarity: 92.078      Percent Identity: 90.815

Match display thresholds for the alignment(s):

| = IDENTITY  
 : = 2  
 . = 1

fbh48003fl.pep x ab021875.pep

```

1  MADSSEGPRAGPGEVAELPGDESGTPGGEAFPLSSLANLFEGEDGSLSPS 50
1  MADPGDGPRAAPGEVAEPPGDESGTSGGEAFPLSSLANLFEGEEGSSSLS 50
51 PADASRPAGPGDGRPNLRMKFQGAFRKGVPNPIDLLESTLYESSVVPGP 100
51 PWMLAALLALAMERPNNLRMKFQGAFRKGVPNPIDLLESTLYESSVVPGP 100
101 KAPMDSLFDYGTyrHHSSDNKRWRKKIIEKQSPKAPAPQPPPIKVF 150
101 KAPMDSLFDYGTyrHHPSDNKRWRKVVVEKQSPKTPAPQPPPIKVF 150
151 RPILFDIVSRGSTADLDGLLPFLTHKKRLTDEEFREPSTGKTCLPKAL 200
151 RPILFDIVSRGSTADLDGLLSFLLTHKKRLTDEEFREPSTGKTCLPKAL 200
201 NLSNGRNDTIPVLLDIAERTGNMREFINSPFRDIYYRGQTALHIAIERR 250
201 NLSNGRNDTLQVLLDIAERTGNMREFINSPFRDIYYRGQTS LHIAIERR 250
251 KHYVELLVAQGADVHAQARGRFFQPKDEGGYFYFGELPLSLAACTNQPH 300
251 KHYVELLVAQGADVHAQARGRFFQPKDEGGYFYFGELPLSLAACTNQPH 300
301 VNYLTENPHKKADMRRQDSRGNTVLHALVAIAD.NTRENTKFVTKMYDL 349
301 VNYLTENPHKKADMRRQDSRGNTVLHALVAIADKHPREHQVCHQDVRPAA 350
350 LLKCARLFPPDSNLEAVLNNDGLSPLMMAAKTGKIGIFQHIIRREVTDED 399
351 SQVVHASSPTATLETVLNNDGLSPLMMAAKTGKIGVFQHIIRREVTDED 400

```

Fig. 4A

|     |   |     |
|-----|---|-----|
| 400 | RHLSRKFKDWAYGPVYSSLYDLSSSLDTCGEEASVLEILVYNSKIENRHEN | 449 |
| 401 | RHLSRKFKDWAYGPVYSSLYDLSSSLDTCGEEVSVLEILVYNSKIENRHEN | 450 |
| 450 | LAVEPINELLRDKWRKFGAVSFYINVVSYLCAMVIFTLTAYYQPLEGTPP  | 499 |
| 451 | LAVEPINELLRDKWRKFGAVSFYINVVSYLCAMVIFTLTAYYQPLEGTPP  | 500 |
| 500 | YPYRTTVDYLRRLAGEVITLFTGVLFFFTNIKDLFMKKCPGVNSLFIDGSF | 549 |
| 501 | YPYRTTVDYLRRLAGEVITLFTGVLFFFTNIKDLFTKKCPGVNSLFVDGSF | 550 |
| 550 | QLLYFIYSVLVIVSAALYLAGIEAYLAVMVFALVLGWMNALYFTRGLKLT  | 599 |
| 551 | QLLYFIYSVLVIVSAALYLAGIEAYLAVMVFALVLGWMNALYFTRGLKLT  | 600 |
| 600 | GTYSIMIQKILFKDLFRFLLVYLLFMIGYASALVSLNPNCANMKVCNEDQ  | 649 |
| 601 | GTYSIMIQKILFKDLFRFLLVYLLFMIGYASALVTLNPNCTNMKVCDEDQ  | 650 |
| 650 | TNCTVPTYPSCRDSETFSTFLDLFKLTIGMGDLEMLSSTKYPVVFILL    | 699 |
| 651 | SNCTVPTYPACRDSETFSAFLDLFKLTIGMGDLEMLSSTKYPVVFILL    | 700 |
| 700 | VTYIILTFVLLLNMALMGETVGQVSKESKHIWKLQWATTILDIERSF     | 749 |
| 701 | VTYIILTFVLLLNMALMGETVGQVSKESKHIWKLQWATTILDIERSF     | 750 |
| 750 | VFLRKAFRSGEMVTVGKSSDGTDDRWCFRVDEVNWSHWNQNLGIINEDP   | 799 |
| 751 | VFLRKAFRSGEMVTVGKSSDGTDDRWCFRVDEVSWSHWNQNLGIINEDP   | 800 |
| 800 | GKNETYQYYGFSHTVGRLLRRDRWSSVPRVVELNKNPNPDEVVPLDSMG   | 849 |
| 801 | GKSEIQYYGFSHTVGRLLRRDRWSSVPRVVELNKNSSADEVVPLDNLG    | 850 |
| 850 | NPRCDGHQQGYPRKWRTDDAPL                              | 871 |
| 851 | NPNCDBGHQQGYAPKWRTDDAPL                             | 872 |

**SUBSTITUTE SHEET (RULE 26)**

11/22

## CLUSTAL W (1.74) multiple sequence alignment

```

Fbh48003FL      TCGACCCACGCGTCCGGGATCTCCCGGCCGCGGGCGCCAGCCGTCCTCGGAGGCTGAGCA
AB021875      -----TCCTGAGGCTGAGAA
                                     ** ***** *

Fbh48003FL      GTGCAGACGGGCCTGGGGCAGGCATGGCGGATTCCAGCGAAGGCCCGCGGGGGCCCG
AB021875      GTACAAACAGATCTGGGTCCAGTATGGCAGATCCTGGTGATGGTCCCCGTGCAGCGCCTG
** ** * * * ***** * * ***** ** * * * * *

Fbh48003FL      GGGAGGTGGCTGAGCTCCCCGGGGATGAGAGTGGCACCCCAGGTGGGGAGGCTTTTCCTC
AB021875      GGGAGGTGGCTGAGCCCCCTGGAGATGAGAGTGGTACCTCTGGTGGGGAGGCCTTCCCC
***** ** * ***** ** * *

Fbh48003FL      TCTCCTCCCTGGCCAATCTGTTTGAGGGGGAGGATGGCTCCCTTTCGCCCTCACCGGCTG
AB021875      TCTCTTCCCTGGCCAATCTGTTTGAGGGGGAGGAAGGCTCCTCTTCTCTTC-CCCGTGG
**** ***** ** * * *

Fbh48003FL      ATGCCAGTCGCCCTGCTGGCCCAGGCGATGGG-CGACCAAATCTGCGCATGAAGTTCCAG
AB021875      ATGCTAGCCGCCCTGCTGGCCCAGGCGATGGAACGTCAAACCTGCGTATGAAGTTCCAG
**** ** ***** ** *****

Fbh48003FL      GGCGCCTTCCGCAAGGGGGTGCCCAACCCCATCGATCTGCTGGAGTCCACCCTATATGAG
AB021875      GGCGCTTTCCGCAAGGGGGTTCCTCAACCCCATGACCTGTTGGAGTCCACCCTGTACGAG
***** ***** ** *****

Fbh48003FL      TCCTCGGTGGTGCTTGGGCCCAAGAAAGCACCCATGGACTCACTGTTTGACTACGGCACC
AB021875      TCCTCAGTAGTGCTTGGGCCCAAGAAAGCGCCATGGATTCCCTGTTGACTACGGCACT
***** ** ***** ** *****

Fbh48003FL      TATCGTCACCACTCCAGTGACAACAAGAGGTGGAGGAAGAAGATCATAGAGAAGCAGCCG
AB021875      TACCGTCACCACCCCAGTGACAACAAGAGATGGAGGAGAAAGGTCGTGGAGAAGCAGCCA
** ***** ** *****

Fbh48003FL      CAGAGCCCCAAAGCCCCCTGCCCCCTCAGCCGCCCCCATCTCTCAAAGTCTTCAACCGGCCT
AB021875      CAGAGCCCCAAACTCTGACCCCCAGCCACCCCCCATCTCTCAAAGTCTTCAATCGGCCC
***** * *****

Fbh48003FL      ATCCTCTTTGACATCGTGTCCCGGGGCTCCACTGCTGACCTGGACGGGCTGCTCCCATTC
AB021875      ATCCTCTTTGACATTGTGTCCCGGGGCTCCACTGCGGACCTAGATGGACTGCTCTCCTTC
***** ***** ** *****

Fbh48003FL      TTGCTGACCCACAAGAAACGCCTAACTGATGAGGAGTTTCGAGAGCCATCTACGGGGAAG
AB021875      TTGTTGACCCACAAGAACGCGCTGACTGATGAGGAGTTCCGGGAGCCGTCCACGGGGAAG
*** ***** ** *****

Fbh48003FL      ACCTGCCTGCCCAAGGCCCTTGCTGAACCTGAGCAATGGCCGCAACGACACCATCCCTGTG
AB021875      ACCTGCCTGCCCAAGGCGCTGCTGAACCTAAGCAACGGGCGCAACGACACCTCCAGGTG
***** ***** ** *****

Fbh48003FL      CTGCTGGACATCGCGGAGCGCACCGGCAACATGAGGGAGTTTCACTTAACTCGCCCTTCCGT
AB021875      TTGCTGGACATTGCGGAGCGCACCGGCAACATGCGTGAATTCATCAACTCGCCCTTCAGA
***** ***** ** *****

Fbh48003FL      GACATCTACTATCGAGGTCAGACAGCCCTGCACATCGCCATTGAGCGTCGCTGCAAACAC
AB021875      GACATCTACTACCGAGGCCAGACATCCCTGCACATTGCCATCGAACGGCGCTGCAAGCAC
***** ***** ** *****

```

Fig. 5A



12/22

Fbh48003FL  
AB021875  
TACGTGGAAC TTCTCGTGGCCCAGGGAGCTGATGTCCACGCCAGGCCCGTGGGCGCTTC  
TACGTGGAGCTGCTGGTGGCCCAGGGAGCCGACGTGCACGCCAGGCCCGCGGCCGCTTC  
\*\*\*\*\* \*\* \*\* \*\*\*\*\* \*\* \*\* \*\*\*\*\*

Fbh48003FL  
AB021875  
TTCCAGCCCAAGGATGAGGGGGCTACTTCTACTTTGGGGAGCTGCCCCGTGCGCTGGCT  
TTCCAGCCCAAGGATGAGGGAGGCTACTTCTACTTTGGGGAGCTGCCCTTGTCCCTGGCA  
\*\*\*\*\* \*\*\*\*\*

Fbh48003FL  
AB021875  
GCCTGCACCAACCAGCCCCACATTGTCAACTACCTGACGGAGAACCCCCACAAGAAGCGG  
GCCTGCACCAACCAGCCGCACATCGTCAACTACCTGACAGAGAACCCTCACAAGAAAGCT  
\*\*\*\*\* \*\*\*\*\*

Fbh48003FL  
AB021875  
GACATGCGGCGCCAGGACTCGCGAGGCAACACAGTGTGCTGCGCTGGTGGCCATTGCT  
GACATGAGGCGACAGGACTCGAGGGGGAACACGGTGTGCTGCACGCGCTGGTGGCCATCGCC  
\*\*\*\*\* \*\*\*\*\*

Fbh48003FL  
AB021875  
GACAA-CACCCGTGAGAACACCAAGTTTGTACCAAGATGTACGACCTGCTGCTGCTCAA  
GACAAACACCCGAGAGAACACCAAGTTTGTACCAAGATGTACGACCTGCTGCTTCTCAA  
\*\*\*\*\* \*\*\*\*\*

Fbh48003FL  
AB021875  
GT-GTGCCCGCTCTTCCCCGACAGCAACCT-GGAGGCCGTGCTCAACAACGACGGCCTC  
GTTGTTCACGCCTCTTCCCCGACAGCAACCTTGGAGACAGTTCTCAACAATGATGGCCTT  
\*\* \* \*\*\*\*\* \*\* \* \*\*\*\*\*

Fbh48003FL  
AB021875  
TCGCCCTCATGATGGCTGCCAAGACGGGCAAGATTGGGATCTTTCAGCACATCATCCGG  
TCGCCTCTCATGATGGCTGCCAAGACAGGCAAGATCGGGGTCTTTCAGCACATCATCCGA  
\*\*\*\*\* \*\*\*\*\*

Fbh48003FL  
AB021875  
CGGGAGGTGACGGATGAGGACACACGGCACCTGTCCCGCAAGTTCAAGGACTGGGCCTAT  
CGTGAGGTGACAGATGAGGACACCCGGCATCTGTCTCGCAAGTTCAAGGACTGGGCCTAT  
\*\* \*\*\*\*\*

Fbh48003FL  
AB021875  
GGGCCAGTGTATTCTCGCTTTATGACCTCTCCTCCCTGGACACGTGTGGGGAAGAGGCC  
GGGCCTGTGTATTCTTCTCTCTACGACCTCTCCTCCCTGGACACATGCGGGGAGGAGGTG  
\*\*\*\*\* \*\*\*\*\*

Fbh48003FL  
AB021875  
TCCGTGCTGGAGATCCTGGTGTACAACAGCAAGATTGAGAACCGCCACGAGATGCTGGCT  
TCCGTGCTGGAGATCCTGGTGTACAACAGCAAGATCGAGAACCGCCATGAGATGCTGGCT  
\*\*\*\*\* \*\*\*\*\*

Fbh48003FL  
AB021875  
GTGGAGCCCATCAATGAACTGCTGCGGGACAAGTGGCGCAAGTTGCGGGCGCTCTCCTTC  
GTAGAGCCCATTAACGAACTGTTGAGAGACAAGTGGCGTAAGTTTGGGGCTGTGTCTTC  
\*\* \*\*\*\*\*

Fbh48003FL  
AB021875  
TACATCAACGTGGTCTCCTACCTGTGTGCCATGGTCATCTTCACTCTCACCGCCTACTAC  
TACATCAACGTGGTCTCCTATCTGTGTGCCATGGTCATCTTCACTCTCACCGCCTACTAT  
\*\*\*\*\* \*\*\*\*\*

Fbh48003FL  
AB021875  
CAGCCGCTGGAGGGCACACCGCCGTACCTTACCGCACCACGGTGGACTACCTGCGGCTG  
CAGCCACTGGAGGGCACGCCACCTACCTTACCGGACCACAGTGGACTACCTGAGGCTG  
\*\*\*\*\* \*\*\*\*\*

Fbh48003FL  
AB021875  
GCTGGCGAGGTCATTACGCTCTTCACTGGGGTCTGTCTTCTTTCACCAACATCAAAGAC  
GCTGGCGAGGTCATCACGCTCTTCAAGGAGTCTGTCTTCTTTCACCAAGATCAAAGAC  
\*\*\*\*\* \*\*\*\*\*

Fbh48003FL  
AB021875  
TTGTTTCATGAAGAAATGCCCTGGAGTGAATTCTCTCTTCAATTGATGGCTCCTTCCAGCTG  
TTGTTTCACGAAGAAATGCCCTGGAGTGAATTCTCTCTTCTCGTCGATGGCTCCTTCCAGTTA  
\*\*\*\*\* \*\*\*\*\*

Fig. 5B

13/22

Fbh48003FL  
AB021875  
CTCTACTTCATCTACTCTGTCTGGTGATCGTCTCAGCAGCCCTCTACCTGGCAGGGATC  
CTCTACTTCATCTACTCTGTCTGGTGGTGTCTCTGCGGCGCTCTACCTGGCTGGGATC  
\*\*\*\*\* \* \*\*\*\*\*

Fbh48003FL  
AB021875  
GAGGCCTACCTGGCCGTGATGGTCTTTGCCCTGGTCCTGGGCTGGATGAATGCCCTTTAC  
GAGGCCTACCTGGCTGTGATGGTCTTTGCCCTGGTCCTGGGCTGGATGAATGCGCTGTAC  
\*\*\*\*\* \*\*

Fbh48003FL  
AB021875  
TTCACCCGTGGGCTGAAGCTGACGGGGACCTATAGCATCATGATCCAGAAGATTCTCTTC  
TTCACGCGCGGGTTGAAGCTGACGGGGACCTACAGCATCATGATTAGAAGATCCTCTTC  
\*\*\*\*\* \*\*

Fbh48003FL  
AB021875  
AAGGACCTTTTTCCGATTCTGCTCGTCTACTTGCTCTTCATGATCGGCTACGCTTCAGCC  
AAAGACCTCTTCCGCTTCTGCTTGTGTACCTGCTCTTCATGATCGGCTATGCCTCAGCC  
\*\* \*\*\*\*\*

Fbh48003FL  
AB021875  
CTGGTCTCCCTCCTGAACCCGTGTGCCAACATGAAGGTGTGCAATGAGGACCAGACCAAC  
CTGGTCACCTCCTGAATCCGTGCACCAACATGAAGGTCTGTGACGAGGACCAGAGCAAC  
\*\*\*\*\* \*\*\*\*\*

Fbh48003FL  
AB021875  
TGCACAGTGCCCACTTACCCCTCGTGCCGTGACAGCGAGACCTTCAGCACCTTCCTCCTG  
TGCACGGTGCCCACTATCCTGCGTGCCGCGACAGCGAGACCTTCAGCGCTTCCTCCTG  
\*\*\*\*\* \*\*\*\*\*

Fbh48003FL  
AB021875  
GACCTGTTTAAGCTGACCATCGGCATGGGCGACCTGGAGATGCTGAGCAGCACCAAGTAC  
GACCTCTTCAAGCTCACCATCGGCATGGGAGACCTGGAGATGCTGAGCAGCGCCAAGTAC  
\*\*\*\*\* \*\*

Fbh48003FL  
AB021875  
CCCGTGGTCTTCATCATCCTGCTGGTGACCTACATCATCCTCACCTTTGTGCTGCTCCTC  
CCCGTGGTCTTCATCCTCCTGCTGGTCACCTACATCATCCTCACCTTCGTGCTCCTGTG  
\*\*\*\*\* \*\*\*\*\*

Fbh48003FL  
AB021875  
AACATGCTCATTGCCCTCATGGGCGAGACAGTGGGCCAGGTCTCCAAGGAGAGCAAGCAC  
AACATGCTTATCGCCCTCATGGGTGAGACCGTGGGCCAGGTGTCCAAGGAGAGCAAGCAC  
\*\*\*\*\* \*\*

Fbh48003FL  
AB021875  
ATCTGGAAGCTGCAGTGGGCCACCAACCATCCTGGACATTGAGCGCTCCTTCCCCGTATTC  
ATCTGGAAGTTGCAGTGGGCCACCAACCATCCTGGACATCGAGCGTTCTTCCCTGTGTTC  
\*\*\*\*\* \*\*\*\*\*

Fbh48003FL  
AB021875  
CTGAGGAAGGCCTTCCGCTCTGGGGAGATGGTCAACGTGGGCAAGAGCTCGGACGGCACT  
CTGAGGAAGGCCTTCCGCTCCGGAGAGATGGTGACTGTGGGCAAGAGCTCAGATGGCACT  
\*\*\*\*\* \*\*

Fbh48003FL  
AB021875  
CCTGACCGCAGGTGGTGCTTCAGGGTGGATGAGGTGAAGTGGTCTCACTGGAACCAGAAC  
CCGACCGCAGGTGGTGCTTCAGGGTGGACGAGGTGAGCTGGTCTCACTGGAACCAGAAC  
\*\* \*\*\*\*\*

Fbh48003FL  
AB021875  
TTGGGCATCATCAACGAGGACCCGGGCAAGAATGAGACCTACCAGTATTATGGCTTCTCG  
TTGGGCATCATTAAACGAGGACCTGGCAAGAGTGAAATCTACCAGTACTATGGCTTCTCC  
\*\*\*\*\* \*\*\*\*\*

Fbh48003FL  
AB021875  
CATACCGTGGGCGCCTCCGACGGGATCGCTGGTCCTCGGTGGTACCCCGGTGGTGGAA  
CACACCGTGGGCGCCTTCGTAGGGATCGTTGGTCCTCGGTGGTGCCCGCGTGTAGTGGAG  
\*\* \*\*\*\*\*

Fbh48003FL  
AB021875  
CTGAACAAGAACTCGAACCCGGACGAGGTGGTGGTGCCTCTGGACAGCATGGGGAACCC  
CTGAACAAGAACTCAAGCGCAGATGAAGTGGTGGTACCCCTGGATAACCTAGGGAACCC  
\*\*\*\*\*

Fig. 5C

14/22

Fbh48003FL  
AB021875  
CGCTGCGATGGCCACCAGCAGGGTTACCCCGCAAGTGGAGGACTGATGACGCCCCGCTC  
AACTGTGACGGCCACCAGCAGGGCTACGCTCCCAACTGGAGGACGGACGATGCCCCACTG  
\*\*\* \*\* \*\*\*\*\* \*\* \* \* \*\*\* \*\*\*\*\* \*\* \*\* \*\*\*\*\* \*\*

Fbh48003FL  
AB021875  
TAGGGACTGCAGCCCAGCCC---CAGCTTCTCTGCCACTCATTCTAGTCCAGCCGCAT  
TAGGGGCCGTGCCAGAGCTCGCACAGATAGTCCAGGCTTGGCCTTCGCTCCACCTACAT  
\*\*\*\*\* \* \* \* \*\*\* \* \*\*\* \* \*\* \* \*\*\* \*\*\*\* \* \*\*\*

Fbh48003FL  
AB021875  
TTCAGCAGTGCCCTTCTGGGGTGTCCCCCACACCCTGCTTTGGCCCCAGAGGCGAGGGAC  
TTAGGCATT---TGTCGGTGTCTTCCACACCCGCTGGGACCTTGGAGGTGAGGGCC  
\*\* \*\*\* \* \* \* \*\*\*\*\* \*\*\*\*\* \* \* \*\* \*\*\*\*\* \*\*

Fbh48003FL  
AB021875  
-CAGTGGAGGTGCCAGGGAGGCCCCAGGACCCTGTGGTCCCC-----TGGC--TC  
TCTGTGGCGACTCTGTGGAGGCCCCAGGACCCTCTGGTCCCCGCCAAGACTTTTGCCCTTC  
\* \*\*\*\* \* \* \*\*\*\*\* \*\*\*\*\* \*\* \*\* \*\*

Fbh48003FL  
AB021875  
TGC---CTCCCCACCCCTGGGGTG-GGGGCTCCCGGCCACCTGTCTTGCTC-----CTAT  
AGCTCTACTCCCCACATGGGGGGGCGGGGCTCCTGGCTACCTGTCTCGCTCGCTCCCCAT  
\*\* \*\*\*\*\* \*\*\*\* \* \*\*\*\*\* \*\* \*\*\*\*\* \*\*\*\* \* \*\*

Fbh48003FL  
AB021875  
GGAGTCACATAAGCCAACGCCAGAGCCC-CTCCAC-----CTCAGGCCCCAGCCCCTG  
GGAGTCACCTAAGCCAGCACAAAGGCCCTCTCCTCGAAAAGGCTCAGGCCCCATCCC--T  
\*\*\*\*\* \*\*\*\*\* \* \* \*\* \*\*\* \*\* \* \*\*\*\*\* \*\*\*\*\* \*\*

Fbh48003FL  
AB021875  
CCTCTCCATTATTTATTTGCTCTGCTCTCAGGAA-----GCGACGTGACCCCTGCCCCA  
CTTGTTATTTATTTATT-GCTCTCCTC--AGGAAAATGGGGTGGCAGGAGTCCACCCGCC  
\* \* \* \*\*\*\*\* \*\*\*\*\* \*\*\* \*\*\*\*\* \* \* \* \*\* \*\* \*\*

Fbh48003FL  
AB021875  
GCTGGAACCTGGCAGAGGCCCTTAGGACCCCGTTCCAAGTGCACTGCCCGGCCAAGCCCCA  
GCTGGAACCTGCCAG-GGCT---GAAGCTCATGCAGGGACGCTGCAGCTCCGA---CCT  
\*\*\*\*\* \*\*\* \*\* \*\* \* \* \* \* \* \*\* \*\* \*\*

Fbh48003FL  
AB021875  
GCCTCAGCCTGCGCCTGAGCTGCATGCGCCACCATTTTGGCAGCGTGGCAGCTTTGCAA  
GCCACAGATCTGACCTG--CTGCAGCCCTGGCTAGTGTGGGT--CTTCTGTACTTTG-AA  
\*\*\* \*\* \*\*\*\* \*\*\*\*\* \* \* \* \* \* \* \* \*\*\*\*\* \*\*

Fbh48003FL  
AB021875  
GGGGCTGGGGCC-CTCGGCGTGGGGCCATGCCTTCTGTGTGTTCTGTAGTGTCTGGGATT  
GAGATTGGGGCCGCTGGNGCTCAANAAANGTTTNTTNTNGNGGNAAAAAAAAAAAAAAAA  
\* \* \*\*\*\*\* \*\* \* \* \* \* \* \* \* \* \*

Fbh48003FL  
AB021875  
TGCCGGTGTCTCAATAAATGGTTATTTCATTGAAAAAAAAAAAAAAAAAAGG-----  
AA  
\*\* \*\*\* \* \* \*\*\*\*\*

Fbh48003FL  
AB021875  
-----  
AAGGGCGGCCGCAAG

Fig. 5D

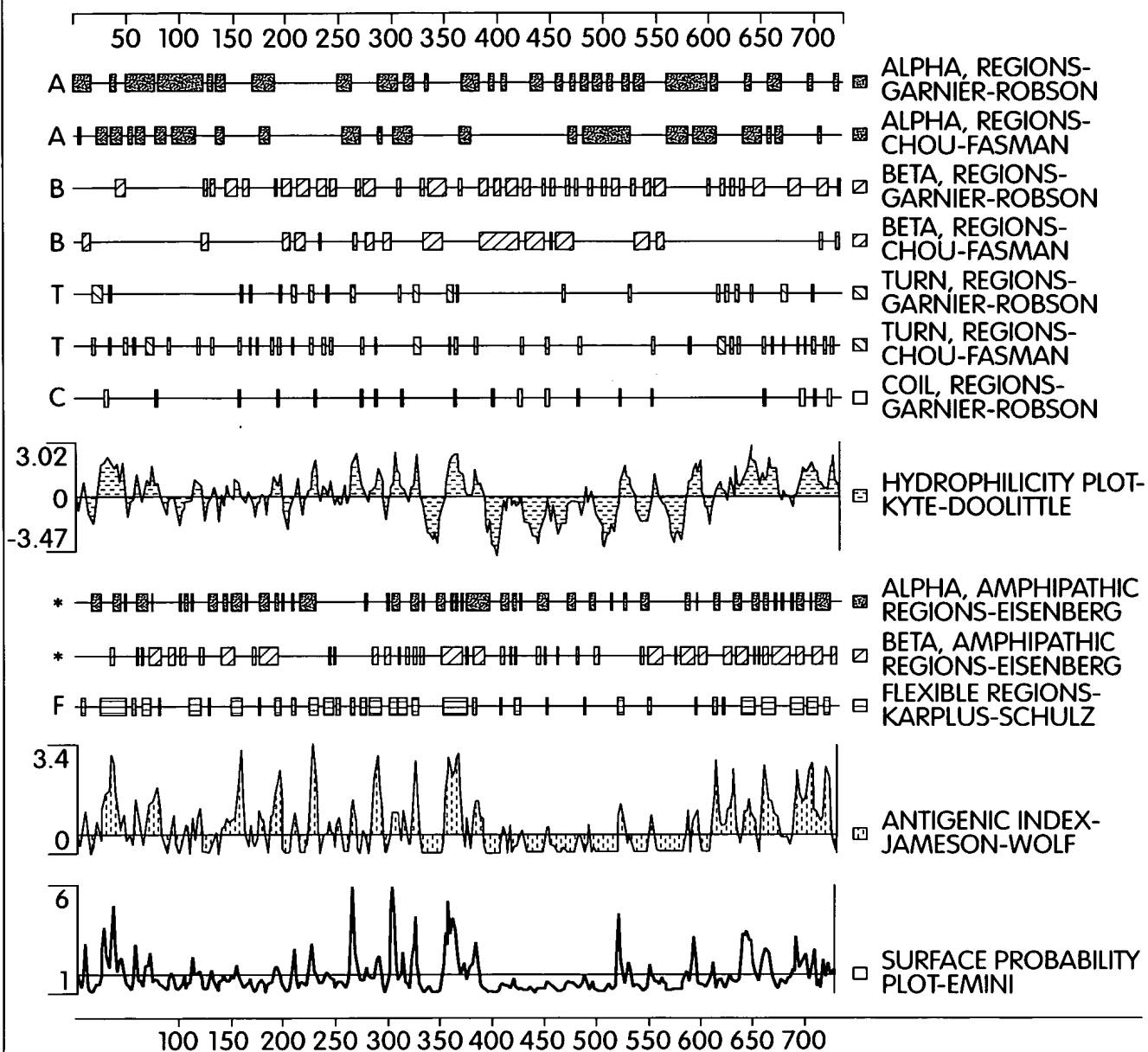


Fig. 6

16/22

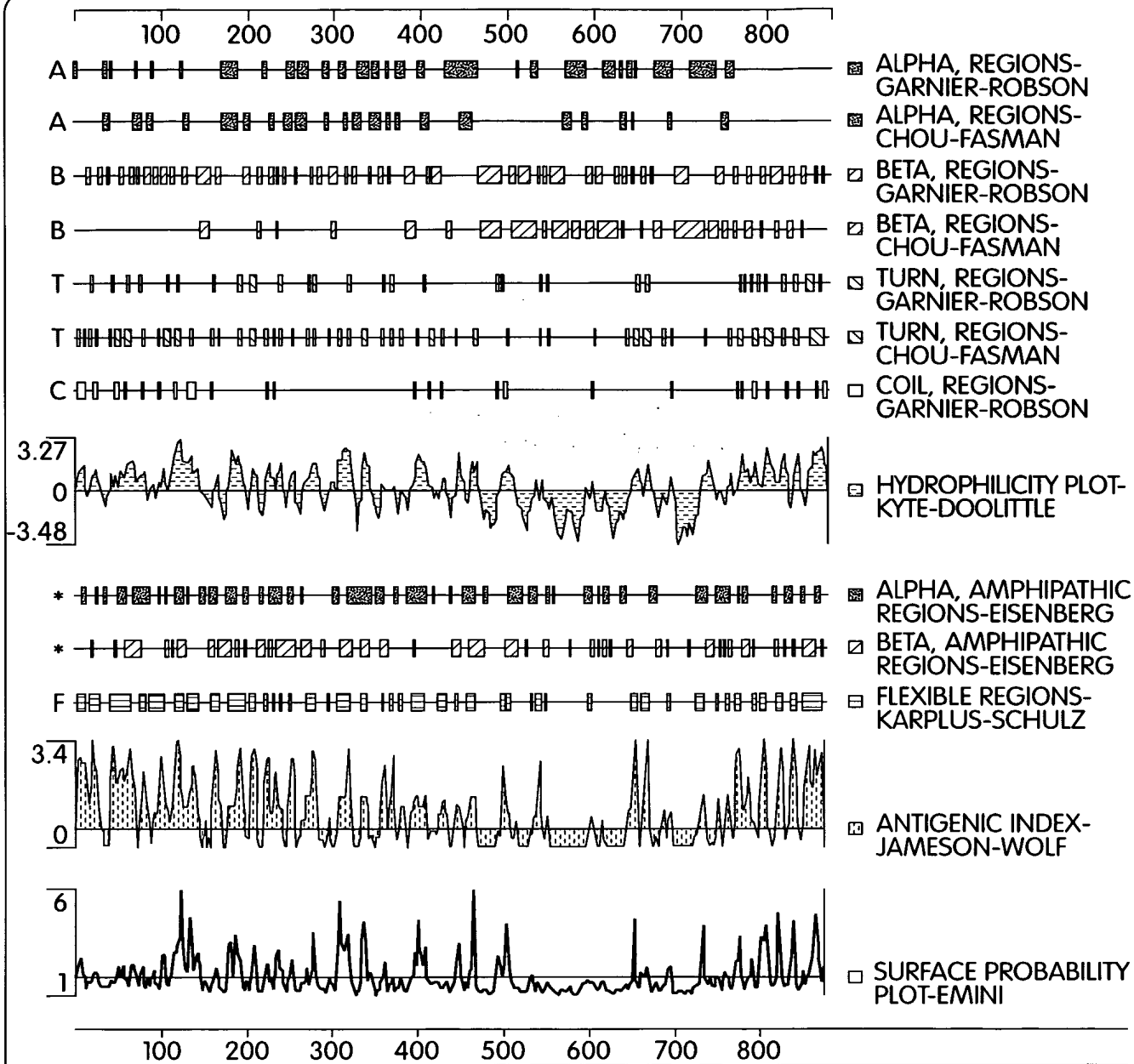


Fig. 7

17/22

Scores for sequence family classification (score includes all domains):

| Model | Description | Score | E-value | N |
|-------|-------------|-------|---------|---|
| ank   | Ank repeat  | 67.2  | 3.5e-16 | 3 |

Parsed for domains

| Model | Domain | seq-f | seq-t | hmm-f | hmm-t   | score | E-value |
|-------|--------|-------|-------|-------|---------|-------|---------|
| ank   | 1/3    | 78    | 108   | ..    | 1 33 () | 16.8  | 0.52    |
| ank   | 2/3    | 116   | 148   | ..    | 1 33 () | 30.7  | 3.4e-05 |
| ank   | 3/3    | 162   | 194   | ..    | 1 33 () | 19.8  | 0.065   |

Alignments of top-scoring domains:

ank: domain 1 of 3, from 78 to 108: score 16.8, E = 0.52

```

      *->nGnTPLHlAarygnvevklLLehGAdvnartk<-*
          G T+LH+Aa y+n+e + +L+e A + + +
18615    78    MGETALHIAALYDNLEAAMVLME--AAPELVFE    108

```

ank: domain 2 of 3, from 116 to 148: score 30.7, E = 3.4e-05

```

      *->nGnTPLHlAarygnvevklLLehGAdvnartk<-*
          G+T+LH+A+ + n+++v++LL + A+v ar
18615    116    EGQTALHIAVVNQNMNLVRALLARRASVSARAT    148

```

ank: domain 3 of 3, from 162 to 194: score 19.8, E = 0.065

```

      *->nGnTPLHlAarygnvevklLLehGAdvnartk<-*
          G PL Aa ++ e+v+lL+ehGAd+ a++
18615    162    FGEHPLSFAACVNSEEIVRLLEHGADIRAQDS    194

```

Fig. 8

18/22

| Model     | Domain | seq-f | seq-t | hmm-f | hmm-t    | score  | E-value |
|-----------|--------|-------|-------|-------|----------|--------|---------|
| ank       | 1/3    | 237   | 269   | ..    | 1 33 ()  | 31.1   | 2.6e-05 |
| ank       | 2/3    | 284   | 319   | ..    | 1 33 ()  | 2.2    | 37      |
| ank       | 3/3    | 369   | 400   | ..    | 1 33 ()  | 12.1   | 2.7     |
| ion_trans | 1/1    | 473   | 718   | ..    | 1 305 () | -133.3 | 0.86    |

## Alignments of top-scoring domains:

ank: domain 1 of 3, from 237 to 269: score 31.1, E = 2.6e-05

```

      *->nGnTPLHLAarygnvevvklLLehGAdvnartk<-*
      +G+T+LH+A + + v+lL+ +GAdv+a+ +
48003 237  RGQTALHIAIERRCKHYVELLVAQGADVHAQAR 269

```

ank: domain 2 of 3, from 284 to 319: score 2.2, E = 37

```

      *->nGnTPLHLAarygnvevvklLLe...hGAdvnartk<-*
      G PL lAa +++ +v++L e++++ Ad ++
48003 284  FGELPLSLAACTNQPHIVNYLTEnphKKADMRRQDS 319

```

ank: domain 3 of 3, from 369 to 400: score 12.1, E = 2.7

```

      *->nGnTPLHLAarygnvevvklLLehGAdvnartk,-*
      +G++PL++Aa++g++ + + ++ + t+
48003 369  DGLSPLMMAAKTGKIGIFQHIIR-REVTDEDTR 400

```

ion\_trans: domain 1 of 1, from 473 to 718: score -133.3, E = 0.86

```

      *->iVsspyFelfilltillNddkvsktialamehpnq...etlndiLdy
      i + +y + +++++t + + +++++ + + +
48003 473  INVVSYLCAmVIFTLTAY---YQP--LEGTPPYPrttVDYLRLAGE 514

      veyvFtgiFtfEmllKmiAlgfkLHkgaYFrng.WNildFvVllsiel
      v ++Ftg+ +f +K + ++ + + ++ F +g++ +l f+ +l i++
48003 515  VITLFTGVLFFFTNIKDLFMKKCPGVNSLFIDGsFQLLYFIYSVLVIVSA 564

      glslinkkaanvggspqqakgslfglkvLRlfrvLRpLKLv.rrapGLrv
      +l l+++ a+ + ++l + ++ R LKL++ ++ ++
48003 565  ALYLAGIEAYLA--VMVF---ALVLGWMNALYFTRGLKLTgTYSIMIQQ 608

      LvqtlInSmkaLgnllLLlflfvfIFAiGmqlF..agkfefdcideste
      ++ +k L+ lL+++lf++ +A++ +l ++ ++ + c++++t
48003 609  IL-----FKDLFRFLLVYLLFMIGYASALVSLLnPCANM-KVCNEDQT- 650

      lfDiiatEpslcgnesyardcpd.gytccrrgWegPnngrtnFdnfpqAfl
      +c+ + y r e F +f +l
48003 651  -----NCTVpTYPSCRDSE-----TFSTF---LL 671

      tlfQvmTgeGWgdvlydtidAagedcdPeseagggicgnnvlmgiiyFisl
      lF + g G ++l t + P v+++i +++
48003 672  DLFKLTIGMGDLEMLSSSTKY-----P-----VVFIIILLVTY 702

      iilgsFltlnLflaVI<-*
      iil ++l+lN+ +a+
48003 703  IILTFVLLLNMLIALM 718

```

Fig. 9

19/22

>141801 p99 (1) 035433\_RAT // VANILLOID RECEPTOR SUBTYPE  
Length = 329

Score = 191 (72.3 bits), Expect = 1.7e-12, P = 1.7e-12  
Identities = 48/130 (36%), Positives = 66/130 (50%)

Query: 100 EAAPELVFEPMTSELYEGQTALHIAVVNQNMXXXXXXXXXXXXXXXXXXXXTGTAFRRSPCNL 159  
++ + V T Y+GQTALHIA+ +NM G F+++  
Sbjct: 184 DSLKQFVNASYTDSYYKGQTALHIAIERRNMTLVTLLEVENGADVQAAANGDFFKKTGRP 243

Query: 160 -IYFGEHPLSFAACVNSEEIVRLLEHG---ADIRAQDSLGNLTVLHILILQPNKT----- 210  
YFGE PLS AAC N IV+ L+++ ADI A+DS+GNTVLH L+ + T  
Sbjct: 244 GFYFGELPLSLAACTNQLAIVKFLQNSWQPADISARDSVGNLTVLHALVEADNTVDNTK 303

Query 211 FACQMYNLLL 220  
F MYN +L  
Sbjct 304 FVTSMYNEIL 313

>2328 P99.2 (23) // PROTEIN CHANNEL CALCIUM RECEPTOR IONIC TRANSMEMBRANE ION  
TRANSPORT ENTRY TRANSIENT  
LENGTH = 272

Score = 96 (38.9 bits), Expect = 1.1e-06, Sum P(2) = 1.1e-06  
Identities = 26/91 (28%), Positives = 46/91 (50%)

Query: 530 LFSTFELFLTIIDGPAN---YNVDLPFMYISIT-----YAAFAIIATLLMLNLLIAMMG 579  
LF TF+ I G A+ Y+ +L + + T + + +I +++LN+LIAMM  
Sbjct: 178 LFETFQSLFWSIFGLAHVDLYSTELSYNHEFTEFVGKVMFGTYNVIMVIVLLNMLIAMMN 237

Query: 580 DTHWRVAHERDELWRAQIVATTVMLEKRLPR 610  
+++ +A D W+ +MLER L +  
Sbjct: 238 NSYQLIADHADVEWKFARAKLWMMLERSLSK 268

Score = 87 (35.7 bits), Expect = 1.1e-06, Sum P(2) = 1.1e-06  
Identities = 35/153 (22%), Positives = 69/153 (45%)

Query: 382 DDIRLVGELVTVIGAIIL-LVEVPDIFRMGVTRFFGQTILGGPFHVLIIITYAFMVLVTM 440  
D RL EL+T+IG I + ++ DI R G +++ + P ++ +L+ +  
Sbjct: 1 DYFRLACELLTIIGCIFFVGFLDFGDIRRQGRKWWN-LLKTFPAKIVFCIANLFLLCI 59

Query 441 VMRL-----ISASGEVVPMSFALVLGWCNVMYFARGFQMLGPFTI-MNQKMIFGDLMR 492  
RL S E + + A V + ++Y + LGP I + +MI D+++  
Sbjct: 60 PFRACKHEFEFSLIAEAL-FAIANVFSYLRLIYIFTANKHLGPLQISLGTRMIV-DIIK 117

Query: 493 FCWLMAVVILGFA---SAFYIIFQTEDPEELGH 522  
F ++ +V+ F+ + Y ++T + H  
Sbjct: 118 FMFIYLLVLFSFSCGLNQLYWYYETSKENKCPH 150

Fig. 10A



20/22

>9056 p99.2 (6) // PROTEIN OLFACTORY CHANNEL B0212.5 T09A12.3 T10B10.7  
 VANILLOID RECEPTOR SUBTYPE F28H7.10  
 Length = 137

Score = 116 (45.9 bits), Expect = 2.6e-06, P = 2.6e-06  
 Identities = 31/102 (30%), Positives = 52/102 (50%)

Query: 257 QHLMQKRKHTQWTYGPLTSTLYDLTEIDS----SGD--EQSLLELIITTKKREARQILDQ 310  
 +HL RK T+W YGP+ + Y L ID+ G+ S + ++ + E ++LD  
 Sbjct: 34 RHL--SRKDTWRYGPFVHCSAYPLNCIDTINEPDGELNSDSAIMTVVYGETVEHLEMLDG 91

Query: 311 TPVKELVSLKWKRYGRPYFCMLGAIYLLYIICFTMCCYIY-RP 351  
 ++ L+ KWK +G+ + M ++ Y CF +C Y RP  
 Sbjct: 92 ELIERLLEDKWKAFGKRLWIMSLGFIFFYCCF-VCAYYLRP 132

>12697 p99.2 (5) // PROTEIN OLFACTORY CHANNEL B0212.5 T09A12.3 F28H7.10  
 T10B10.7  
 Length = 200

Score = 92 (37.4 bits), Expect = 0.072, P = 0.070  
 Identities = 28/87 (32%), Positives = 43/87 (49%)

Query: 51 LAAKDNDVQALNKLLKYEDC--KVHQRGAMGETALHIAALYDNL---EAAMVLMEAAPEL 105  
 L + V KY + K+ +RG+MGE +HI L ++ E A L+ P+L  
 Sbjct: 56 LGLSEESVDMQQSEFKYREL VWKLDERGSMGENLIHICLLRNSQIHNEIARRLLNRFPKL 115

Query: 106 VFEPMTSELYE---GQTALHIAVVNQN 129  
 V + SE Y G + LH+A+VN +  
 Sbjct: 116 VNDIYLSSEYYASVGLSPLHLAIVNDD 142

Fig. 10B

21/22

Length = 329

Score = 618 (22.6 bits), Expect = 1.0e-60, P = 1.0e-6

Identities 24/220 (56%), Positives = 162/220 (73%)

Query: 147 KVFNRPIFLDIVSRGSTADLDGLLPFLLTTHKKRLTDEEFREPSTGKTCLPKALLNLSNGR 206  
 ++++R +FD V++ + +L+ LLPFL KKRLTD EF++P TGKTCL KA+LNL NG+

Sbjct: 110 RLYDRRSIFDAVAQSNCOELESLLPFLQRSKKRLTDSEFKDPETGKTCLLKAMLNLHNGQ 169

Query: 207 NDTIPVLLDIAERTGNMREFINSPPFRDIYYRGQTALHIAIERRCKHYVELLVAQGADVHA 266  
 NDTI +LLD+A +T ++++F+N+ + D YY+GQTALHIAIERR V LLV GADV A

Sbjct: 170 NDTIALLLDVARKTDSLKQFVNASYTDSYKGGQTALHIAIERRNMTLVTLVENGADVQA 229

Query: 267 QARGRFFQP-KDEGGYFYFGELPLSLAACTNQPHIVNYLTENPHKKADMRRQDSRGNTVL 325  
 A G FF+ K G FYFGELPLSLAACTNQ IV +L +N + AD+ +DS GNTVL

Sbjct: 230 AANGDFFKKTGRPG-FYFGELPLSLAACTNQLAIVKFLQNSWQPADISARDSVGNTVL 288

Query: 326 HALVAIADNTRENTKFKVTMYDLLLLKCARLFPDSNLEAV 365  
 HALV +ADNT +NTKFT MY+ +L+ A+L P LE +

Sbjct: 289 HALVEVADNTVDNTKFKVTSMYNEILILGAKLHPTLKLEEI 328

&gt;145518 p99.2 (1) 035433\_RAT // VANILLOID RECEPTOR SUBTYPE 1

Length = 124

Score = 212 (79.7 bits), Expect = 1.8e-16, P = 1.8e-16

Identities = 42/79 (53%), Positives = 53/79 (67%)

Query: 752 LRKAFRSGEMVTVGKSSDGTDPDRWCFRVDEVNWSHWNQNLGIINEDPGKNE-TYQYYGF 810  
 +RKAFRSG+++ VG + DG D RWCFRVDEVNW+ WN N+GIINEDPG E + F

Sbjct: 2 MRKAFRSGKLLQVGFTPDGKDDYRWCFRVDEVNWTWNINVGIINEDPGNCEGVKRTLSF 61

Query: 811 SHTVGRLRRDRWSS--VVP 827  
 S GR+ W + +VP

Sbjct: 62 SLRSGRVSGRNWKNFALVP 80

>9056 p99.2 (6) // PROTEIN OLFACTORY CHANNEL B0212.5 T09A12.3 T10B10.7  
 VANILLOID RECEPTOR SUBTYPE F28H7.10

Length = 137

Score = 196 (74.1 bits), Expect = 9.3e-15, P = 9.3e-15

Identities = 48/138 (34%), Positives = 69/138 (50%)

Query: 367 NNDGLSPLMMAAKTGKIGIFQHIIRREVTDIEDTRHLSRKFKDWAYGPVXXXXXXXXXXXXXT 426  
 N+ GL+PL +AAK GK IF I+ E+ + + RHLSRK +W YGPV T

Sbjct: 1 NHKGLTPLTLAAKLGKKHIFDEILECEIHEPECRHLSRKDTEWRYGPVHCSAYPLNCIDT 60

Query: 427 CGE-----EASVLEILVYNSKIENRHEMLAVEPINELLRDKWRKFGAVSFYINVVSYLC 480  
 E S + +VY +E+ EML E I LL DKW+ FG + ++++ ++

Sbjct: 61 INEPDGELNSDSAIMTVVYGETVEHL-EMLDGELIERLLEDKWKAFGKRLWIMSLLGFI 119

Query: 481 AMVIFTLTAYYQPLEGTP 498  
 F Y +P E P

Sbjct: 120 YYCCFVCAYYLRPTLELP 137

Fig. 11A

22/22

>2328 p99.2 (23) // PROTEIN CHANNEL CALCIUM RECEPTOR IONIC TRANSMEMBRANE ION  
TRANSPORT ENTRY TRANSIENT  
Length = 272

Score = 97 (39.2 bits), Expect = 0.050, P = 0.048  
Identities = 52/209 (24%), Positives = 89/209 (42%)

Query: 561 IVSAAALY-LAGIEAYLAVMVFALVLGWMNALYFTRGLKLTGTYSIMIQILFKDLFRFLL 619  
+++ AL+ +A + +YL ++ + L + G + MI I+ K +F +LL  
Sbjct: 73 LIAEALFAIANVFSYLRLIYIFTANKHLGPLQISLGTR-----MIVDII-KFMFIYLL 124

Query: 620 VYLLEFMIG---YASALVSLNPCANMK--VCNEDQTNCTVPTYPSCRD--SETFSTFLL 671  
V F G Y S N C + + + N D + D S F TF  
Sbjct: 125 VLFSFSCGLNQLYWYYETSKENKCPHCYHLHNADYDCVGISCEQQSNDTFSNLFETFQ- 183

Query: 672 DLFKLTIGMGDLEMLSST-KYPVVFI-----ILLVTYIILTFVLLLNMLIALMGETVGQV 725  
LF G+ +++ S+ Y F ++ TY ++ ++LLNMLIA+M + +  
Sbjct: 184 SLFWSIFGLAHVDLYSTELSYNHEFTEFVGKVMFGTYNVIMVIVLLNMLIAMMNNSYQLI 243

Query: 726 SKESKHIWKLQWATTILDIERSSFPVFLRK 754  
+ + WK A + +ERS RK  
Sbjct: 244 ADHADVEWKFARAKLWMMLERSLSKKERK 272

Score = 88 (36.0 bots), Expect = 0.49, P = 0.39  
Identities = 31/137 (22%), Positives = 64/137 (46%)

Query: 507 DYLRILAGEVITLFTGVLF--FTNIKDLFMKKCPGVNSLFIDGSFQLLYFIYSVLVIVSA 564  
DY RLA E++T+ G +FF F + D+ + +L +++++ I ++ +++  
Sbjct: 1 DYFRLACELLTII-GCIEFFVGFLDFGDIRRQGRKWWNLLKTTPAKIVFCIANLFLICI 59

Query: 565 ALYLAGIEAY----LAVMVFAL--VLGWMNALYFTRGLKLTGTYSIMIQILFKDLFRFL 618  
LA + +A +FA+ V ++ +Y K G I + + D+ +F+  
Sbjct: 60 PFRACKHEFEFSLIAEALFAIANVFSYLRLIYIFTANKHLGPLQISLGTRMIVDIIKFM 119

Query: 619 LVYLLFMIGYASALVSL 635  
+YLL + ++ L L  
Sbjct: 120 FIYLLVLFSFSCGLNQL 136

Fig. 11B

## SEQUENCE LISTING

<110> Millennium Pharmaceuticals, Inc., et al.

<120> 18615 AND 48003, NOVEL HUMAN ION CHANNELS AND  
USES THEREFOR

<130> MNI-095CPPC

<140>

<141>

<150> USSN 09/525,420

<151> 2000-03-15

<160> 6

<170> PatentIn Ver. 2.0

<210> 1

<211> 3026

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (278)..(2455)

<220>

<223> All occurrences of n = any nucleotide

<400> 1

```

gatcnagnnt ntcentcttg tnaacctcgt acgcncgcg gaattcccgg gtcgaccac 60
gcgtccggcc aagtgtaca aactcacagc cctctccaaa ctggctgggg ctgctgggag 120
actcccaagg aactcgtcag gaaggcagga gacaggagac gggacctcta caggagagac 180
gtgggccggc cttggggggg ctgatgtggc cccaaggctg agtcccgta gggctctggcc 240
tcggcctcag gcccccaagg agccggccct acacccc atg ggt ttg tca ctg ccc 295
                               Met Gly Leu Ser Leu Pro
                               1               5

aag gag aaa ggg cta att ctc tgc cta tgg agc aag ttc tgc aga tgg 343
Lys Glu Lys Gly Leu Ile Leu Cys Leu Trp Ser Lys Phe Cys Arg Trp
              10                15                20

ttc cag aga cgg gag tcc tgg gcc cag agc cga gat gag cag aac ctg 391
Phe Gln Arg Arg Glu Ser Trp Ala Gln Ser Arg Asp Glu Gln Asn Leu
              25                30                35

ctg cag cag aag agg atc tgg gag tct cct ctc ctt cta gct gcc aaa 439
Leu Gln Gln Lys Arg Ile Trp Glu Ser Pro Leu Leu Leu Ala Ala Lys
              40                45                50

gat aat gat gtc cag gcc ctg aac aag ttg ctc aag tat gag gat tgc 487
Asp Asn Asp Val Gln Ala Leu Asn Lys Leu Leu Lys Tyr Glu Asp Cys
              55                60                65                70

aag gtg cac cag aga gga gcc atg ggg gaa aca gcg cta cac ata gca 535

```

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |      |  |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|--|
| Lys | Val | His | Gln | Arg | Gly | Ala | Met | Gly | Glu | Thr | Ala | Leu | His | Ile | Ala |      |  |
|     |     |     |     | 75  |     |     |     |     | 80  |     |     |     |     | 85  |     |      |  |
| gcc | ctc | tat | gac | aac | ctg | gag | gcc | gcc | atg | gtg | ctg | atg | gag | gct | gcc | 583  |  |
| Ala | Leu | Tyr | Asp | Asn | Leu | Glu | Ala | Ala | Met | Val | Leu | Met | Glu | Ala | Ala |      |  |
|     |     |     | 90  |     |     |     |     | 95  |     |     |     |     | 100 |     |     |      |  |
| ccg | gag | ctg | gtc | ttt | gag | ccc | atg | aca | tct | gag | ctc | tat | gag | ggg | cag | 631  |  |
| Pro | Glu | Leu | Val | Phe | Glu | Pro | Met | Thr | Ser | Glu | Leu | Tyr | Glu | Gly | Gln |      |  |
|     |     | 105 |     |     |     |     | 110 |     |     |     |     | 115 |     |     |     |      |  |
| act | gca | ctg | cac | atc | gct | gtt | gtg | aac | cag | aac | atg | aac | ctg | gtg | cga | 679  |  |
| Thr | Ala | Leu | His | Ile | Ala | Val | Val | Asn | Gln | Asn | Met | Asn | Leu | Val | Arg |      |  |
|     | 120 |     |     |     |     | 125 |     |     |     | 130 |     |     |     |     |     |      |  |
| gcc | ctg | ctt | gcc | cgc | agg | gcc | agt | gtc | tct | gcc | aga | gcc | aca | ggc | act | 727  |  |
| Ala | Leu | Leu | Ala | Arg | Arg | Ala | Ser | Val | Ser | Ala | Arg | Ala | Thr | Gly | Thr |      |  |
| 135 |     |     |     |     | 140 |     |     |     |     | 145 |     |     |     |     | 150 |      |  |
| gcc | ttc | cgc | cgt | agt | ccc | tgc | aac | ctc | atc | tac | ttt | ggg | gag | cac | cct | 775  |  |
| Ala | Phe | Arg | Arg | Ser | Pro | Cys | Asn | Leu | Ile | Tyr | Phe | Gly | Glu | His | Pro |      |  |
|     |     |     |     | 155 |     |     |     |     | 160 |     |     |     |     | 165 |     |      |  |
| ttg | tcc | ttt | gct | gcc | tgt | gtg | aac | agt | gag | gag | atc | gtg | cgg | ctg | ctc | 823  |  |
| Leu | Ser | Phe | Ala | Ala | Cys | Val | Asn | Ser | Glu | Glu | Ile | Val | Arg | Leu | Leu |      |  |
|     |     |     | 170 |     |     |     |     | 175 |     |     |     |     | 180 |     |     |      |  |
| att | gag | cat | gga | gct | gac | atc | cgg | gcc | cag | gac | tcc | ctg | gga | aac | aca | 871  |  |
| Ile | Glu | His | Gly | Ala | Asp | Ile | Arg | Ala | Gln | Asp | Ser | Leu | Gly | Asn | Thr |      |  |
|     |     | 185 |     |     |     |     | 190 |     |     |     |     | 195 |     |     |     |      |  |
| gtg | tta | cac | atc | ctc | atc | ctc | cag | ccc | aac | aaa | acc | ttt | gcc | tgc | cag | 919  |  |
| Val | Leu | His | Ile | Leu | Ile | Leu | Gln | Pro | Asn | Lys | Thr | Phe | Ala | Cys | Gln |      |  |
|     | 200 |     |     |     |     | 205 |     |     |     |     | 210 |     |     |     |     |      |  |
| atg | tac | aac | ctg | ttg | ctg | tcc | tac | gac | aga | cat | ggg | gac | cac | ctg | cag | 967  |  |
| Met | Tyr | Asn | Leu | Leu | Leu | Ser | Tyr | Asp | Arg | His | Gly | Asp | His | Leu | Gln |      |  |
| 215 |     |     |     |     | 220 |     |     |     |     | 225 |     |     |     |     | 230 |      |  |
| ccc | ctg | gac | ctc | gtg | ccc | aat | cac | cag | ggg | ctc | acc | cct | ttc | aag | ctg | 1015 |  |
| Pro | Leu | Asp | Leu | Val | Pro | Asn | His | Gln | Gly | Leu | Thr | Pro | Phe | Lys | Leu |      |  |
|     |     |     |     | 235 |     |     |     |     | 240 |     |     |     |     | 245 |     |      |  |
| gct | gga | gtg | gag | ggg | aac | act | gtg | atg | ttt | cag | cac | ctg | atg | cag | aag | 1063 |  |
| Ala | Gly | Val | Glu | Gly | Asn | Thr | Val | Met | Phe | Gln | His | Leu | Met | Gln | Lys |      |  |
|     |     |     | 250 |     |     |     |     | 255 |     |     |     |     | 260 |     |     |      |  |
| cgg | aag | cac | acc | cag | tgg | acg | tat | gga | cca | ctg | acc | tcg | act | ctc | tat | 1111 |  |
| Arg | Lys | His | Thr | Gln | Trp | Thr | Tyr | Gly | Pro | Leu | Thr | Ser | Thr | Leu | Tyr |      |  |
|     |     | 265 |     |     |     |     | 270 |     |     |     |     | 275 |     |     |     |      |  |
| gac | ctc | aca | gag | atc | gac | tcc | tca | ggg | gat | gag | cag | tcc | ctg | ctg | gaa | 1159 |  |
| Asp | Leu | Thr | Glu | Ile | Asp | Ser | Ser | Gly | Asp | Glu | Gln | Ser | Leu | Leu | Glu |      |  |
|     | 280 |     |     |     |     | 285 |     |     |     |     | 290 |     |     |     |     |      |  |
| ctt | atc | atc | acc | acc | aag | aag | cgg | gag | gct | cgc | cag | atc | ctg | gac | cag | 1207 |  |
| Leu | Ile | Ile | Thr | Thr | Lys | Lys | Arg | Glu | Ala | Arg | Gln | Ile | Leu | Asp | Gln |      |  |
| 295 |     |     |     |     | 300 |     |     |     |     | 305 |     |     |     |     | 310 |      |  |
| acg | ccg | gtg | aag | gag | ctg | gtg | agc | ctc | aag | tgg | aag | cgg | tac | ggg | cgg | 1255 |  |
| Thr | Pro | Val | Lys | Glu | Leu | Val | Ser | Leu | Lys | Trp | Lys | Arg | Tyr | Gly | Arg |      |  |

| 315 |     |     |     | 320 |     |     |     | 325 |     |     |     |     |     |     |     |      |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|
| ccg | tac | ttc | tgc | atg | ctg | ggt | gcc | ata | tat | ctg | ctg | tac | atc | atc | tgc | 1303 |
| Pro | Tyr | Phe | Cys | Met | Leu | Gly | Ala | Ile | Tyr | Leu | Leu | Tyr | Ile | Ile | Cys |      |
| 330 |     |     |     | 335 |     |     |     | 340 |     |     |     |     |     |     |     |      |
| ttc | acc | atg | tgc | tgc | atc | tac | cgc | ccc | ctc | aag | ccc | agg | acc | aat | aac | 1351 |
| Phe | Thr | Met | Cys | Cys | Ile | Tyr | Arg | Pro | Leu | Lys | Pro | Arg | Thr | Asn | Asn |      |
| 345 |     |     |     | 350 |     |     |     | 355 |     |     |     |     |     |     |     |      |
| cgc | acg | agc | ccc | cgg | gac | aac | acc | ctc | tta | cag | cag | aag | cta | ctt | cag | 1399 |
| Arg | Thr | Ser | Pro | Arg | Asp | Asn | Thr | Leu | Leu | Gln | Gln | Lys | Leu | Leu | Gln |      |
| 360 |     |     |     | 365 |     |     |     | 370 |     |     |     |     |     |     |     |      |
| gaa | gcc | tac | atg | acc | cct | aag | gac | gat | atc | cgg | ctg | gtc | ggg | gag | ctg | 1447 |
| Glu | Ala | Tyr | Met | Thr | Pro | Lys | Asp | Asp | Ile | Arg | Leu | Val | Gly | Glu | Leu |      |
| 375 |     |     |     | 380 |     |     |     | 385 |     |     |     | 390 |     |     |     |      |
| gtg | act | gtc | att | ggg | gct | atc | atc | atc | ctg | ctg | gta | gag | gtt | cca | gac | 1495 |
| Val | Thr | Val | Ile | Gly | Ala | Ile | Ile | Ile | Leu | Leu | Val | Glu | Val | Pro | Asp |      |
| 395 |     |     |     | 400 |     |     |     | 405 |     |     |     |     |     |     |     |      |
| atc | ttc | aga | atg | ggg | gtc | act | cgc | ttc | ttt | gga | cag | acc | atc | ctt | ggg | 1543 |
| Ile | Phe | Arg | Met | Gly | Val | Thr | Arg | Phe | Phe | Gly | Gln | Thr | Ile | Leu | Gly |      |
| 410 |     |     |     | 415 |     |     |     | 420 |     |     |     |     |     |     |     |      |
| ggc | cca | ttc | cat | gtc | ctc | atc | atc | acc | tat | gcc | ttc | atg | gtg | ctg | gtg | 1591 |
| Gly | Pro | Phe | His | Val | Leu | Ile | Ile | Thr | Tyr | Ala | Phe | Met | Val | Leu | Val |      |
| 425 |     |     |     | 430 |     |     |     | 435 |     |     |     |     |     |     |     |      |
| acc | atg | gtg | atg | cgg | ctc | atc | agt | gcc | agc | ggg | gag | gtg | gta | ccc | atg | 1639 |
| Thr | Met | Val | Met | Arg | Leu | Ile | Ser | Ala | Ser | Gly | Glu | Val | Val | Pro | Met |      |
| 440 |     |     |     | 445 |     |     |     | 450 |     |     |     |     |     |     |     |      |
| tcc | ttt | gca | ctc | gtg | ctg | ggc | tgg | tgc | aac | gtc | atg | tac | ttt | gcc | cga | 1687 |
| Ser | Phe | Ala | Leu | Val | Leu | Gly | Trp | Cys | Asn | Val | Met | Tyr | Phe | Ala | Arg |      |
| 455 |     |     |     | 460 |     |     |     | 465 |     |     |     | 470 |     |     |     |      |
| gga | ttc | cag | atg | cta | ggc | ccc | ttc | acc | atc | atg | aat | cag | aag | atg | att | 1735 |
| Gly | Phe | Gln | Met | Leu | Gly | Pro | Phe | Thr | Ile | Met | Asn | Gln | Lys | Met | Ile |      |
| 475 |     |     |     | 480 |     |     |     | 485 |     |     |     |     |     |     |     |      |
| ttt | ggc | gac | ctg | atg | cga | ttc | tgc | tgg | ctg | atg | gct | gtg | gtc | atc | ctg | 1783 |
| Phe | Gly | Asp | Leu | Met | Arg | Phe | Cys | Trp | Leu | Met | Ala | Val | Val | Ile | Leu |      |
| 490 |     |     |     | 495 |     |     |     | 500 |     |     |     |     |     |     |     |      |
| ggc | ttt | gct | tca | gcc | ttc | tat | atc | atc | ttc | cag | aca | gag | gac | ccc | gag | 1831 |
| Gly | Phe | Ala | Ser | Ala | Phe | Tyr | Ile | Ile | Phe | Gln | Thr | Glu | Asp | Pro | Glu |      |
| 505 |     |     |     | 510 |     |     |     | 515 |     |     |     |     |     |     |     |      |
| gag | cta | ggc | cac | ttc | tac | gac | tac | ccc | atg | gcc | ctg | ttc | agc | acc | ttc | 1879 |
| Glu | Leu | Gly | His | Phe | Tyr | Asp | Tyr | Pro | Met | Ala | Leu | Phe | Ser | Thr | Phe |      |
| 520 |     |     |     | 525 |     |     |     | 530 |     |     |     |     |     |     |     |      |
| gag | ctg | ttc | ctt | acc | atc | atc | gat | ggc | cca | gcc | aac | tac | aac | gtg | gac | 1927 |
| Glu | Leu | Phe | Leu | Thr | Ile | Ile | Asp | Gly | Pro | Ala | Asn | Tyr | Asn | Val | Asp |      |
| 535 |     |     |     | 540 |     |     |     | 545 |     |     |     | 550 |     |     |     |      |
| ctg | ccc | ttc | atg | tac | agc | atc | acc | tat | gct | gcc | ttt | gcc | atc | atc | gcc | 1975 |
| Leu | Pro | Phe | Met | Tyr | Ser | Ile | Thr | Tyr | Ala | Ala | Phe | Ala | Ile | Ile | Ala |      |
| 555 |     |     |     | 560 |     |     |     | 565 |     |     |     |     |     |     |     |      |

|  |      |
|--|------|
| aca ctg ctc atg ctc aac ctc ctc att gcc atg atg ggc gac act cac    | 2023 |
| Thr Leu Leu Met Leu Asn Leu Leu Ile Ala Met Met Gly Asp Thr His    |      |
| 570 575 580  |      |
| tgg cga gtg gcc cat gag cgg gat gag ctg tgg agg gcc cag att gtg    | 2071 |
| Trp Arg Val Ala His Glu Arg Asp Glu Leu Trp Arg Ala Gln Ile Val    |      |
| 585 590 595  |      |
| gcc acc acg gtg atg ctg gag cgg aag ctg cct cgc tgc ctg tgg cct    | 2119 |
| Ala Thr Thr Val Met Leu Glu Arg Lys Leu Pro Arg Cys Leu Trp Pro    |      |
| 600 605 610  |      |
| cgc tcc ggg atc tgc gga cgg gag tat gcc ctg gga gac cgc tgg ttc    | 2167 |
| Arg Ser Gly Ile Cys Gly Arg Glu Tyr Gly Leu Gly Asp Arg Trp Phe    |      |
| 615 620 625 630  |      |
| ctg cgg gtg gaa gac agg caa gat ctc aac cgg cag cgg atc caa cgc    | 2215 |
| Leu Arg Val Glu Asp Arg Gln Asp Leu Asn Arg Gln Arg Ile Gln Arg    |      |
| 635 640 645  |      |
| tac gca cag gcc ttc cac acc cgg ggc tct gag gat ttg gac aaa gac    | 2263 |
| Tyr Ala Gln Ala Phe His Thr Arg Gly Ser Glu Asp Leu Asp Lys Asp    |      |
| 650 655 660  |      |
| tca gtg gaa aaa cta gag ctg ggc tgt ccc ttc agc ccc cac ctg tcc    | 2311 |
| Ser Val Glu Lys Leu Glu Leu Gly Cys Pro Phe Ser Pro His Leu Ser    |      |
| 665 670 675  |      |
| ctt cct atg ccc tca gtg tct cga agt acc tcc cgc agc agt gcc aat    | 2359 |
| Leu Pro Met Pro Ser Val Ser Arg Ser Thr Ser Arg Ser Ser Ala Asn    |      |
| 680 685 690  |      |
| tgg gaa agg ctt cgg caa ggg acc ctg agg aga gac ctg cgt ggg ata    | 2407 |
| Trp Glu Arg Leu Arg Gln Gly Thr Leu Arg Arg Asp Leu Arg Gly Ile    |      |
| 695 700 705 710  |      |
| atc aac agg ggt ctg gag gac ggg gag agc tgg gaa tat cag atc tga    | 2455 |
| Ile Asn Arg Gly Leu Glu Asp Gly Glu Ser Trp Glu Tyr Gln Ile        |      |
| 715 720 725  |      |
| ctgcgtgttc tcaattcgtc tccctggaact tgctctcatt ttcttgggtg catcaaacaa | 2515 |
| aacaaaaacc aaacacccag aggtctcatc tcccaggccc caggggagaa agaggagtag  | 2575 |
| catgaacgcc aaggaatgta cgttgagaat cactgctcca ggcttgcatt actccttcag  | 2635 |
| ctctggggca gaggaagccc agcccaagca cggggctggc agggcgtgag gaactctcct  | 2695 |
| gtggcctgct catcaccott ccgacaggag cactgcatgt cagagcactt taaaaacagg  | 2755 |
| ccagcctgct tgggcgctcg gtctccaccc cagggtcata agtggggaga gagcccttcc  | 2815 |
| cagggcaccc aggcaggtgc agggaaagtgc agagcttgtg gaaagcgtgt gagtgaggga | 2875 |
| gacaggaacg gctctggggg tgggaagtgg ggctaggtct tgccaactcc atcttcaata  | 2935 |
| aagtcgtttt cggaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaag gaksnagct   | 2995 |
| tggacgttgc caatcgtcnt ggtgngtttt c                                 | 3026 |

&lt;210&gt; 2

&lt;211&gt; 725

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 2

```

Met Gly Leu Ser Leu Pro Lys Glu Lys Gly Leu Ile Leu Cys Leu Trp
 1           5           10           15

Ser Lys Phe Cys Arg Trp Phe Gln Arg Arg Glu Ser Trp Ala Gln Ser
          20           25           30

Arg Asp Glu Gln Asn Leu Leu Gln Gln Lys Arg Ile Trp Glu Ser Pro
      35           40           45

Leu Leu Leu Ala Ala Lys Asp Asn Asp Val Gln Ala Leu Asn Lys Leu
 50           55           60

Leu Lys Tyr Glu Asp Cys Lys Val His Gln Arg Gly Ala Met Gly Glu
 65           70           75           80

Thr Ala Leu His Ile Ala Ala Leu Tyr Asp Asn Leu Glu Ala Ala Met
          85           90           95

Val Leu Met Glu Ala Ala Pro Glu Leu Val Phe Glu Pro Met Thr Ser
      100           105           110

Glu Leu Tyr Glu Gly Gln Thr Ala Leu His Ile Ala Val Val Asn Gln
 115           120           125

Asn Met Asn Leu Val Arg Ala Leu Leu Ala Arg Arg Ala Ser Val Ser
 130           135           140

Ala Arg Ala Thr Gly Thr Ala Phe Arg Arg Ser Pro Cys Asn Leu Ile
 145           150           155           160

Tyr Phe Gly Glu His Pro Leu Ser Phe Ala Ala Cys Val Asn Ser Glu
          165           170           175

Glu Ile Val Arg Leu Leu Ile Glu His Gly Ala Asp Ile Arg Ala Gln
          180           185           190

Asp Ser Leu Gly Asn Thr Val Leu His Ile Leu Ile Leu Gln Pro Asn
 195           200           205

Lys Thr Phe Ala Cys Gln Met Tyr Asn Leu Leu Leu Ser Tyr Asp Arg
 210           215           220

His Gly Asp His Leu Gln Pro Leu Asp Leu Val Pro Asn His Gln Gly
 225           230           235           240

Leu Thr Pro Phe Lys Leu Ala Gly Val Glu Gly Asn Thr Val Met Phe
          245           250           255

Gln His Leu Met Gln Lys Arg Lys His Thr Gln Trp Thr Tyr Gly Pro
          260           265           270

Leu Thr Ser Thr Leu Tyr Asp Leu Thr Glu Ile Asp Ser Ser Gly Asp
 275           280           285

```



|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  |  |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|--|
| Glu | Gln | Ser | Leu | Leu | Glu | Leu | Ile | Ile | Thr | Thr | Lys | Lys | Arg | Glu | Ala |  |  |
|     | 290 |     |     |     |     | 295 |     |     |     |     | 300 |     |     |     |     |  |  |
| Arg | Gln | Ile | Leu | Asp | Gln | Thr | Pro | Val | Lys | Glu | Leu | Val | Ser | Leu | Lys |  |  |
| 305 |     |     |     |     | 310 |     |     |     |     | 315 |     |     |     |     | 320 |  |  |
| Trp | Lys | Arg | Tyr | Gly | Arg | Pro | Tyr | Phe | Cys | Met | Leu | Gly | Ala | Ile | Tyr |  |  |
|     |     |     |     | 325 |     |     |     |     | 330 |     |     |     |     | 335 |     |  |  |
| Leu | Leu | Tyr | Ile | Ile | Cys | Phe | Thr | Met | Cys | Cys | Ile | Tyr | Arg | Pro | Leu |  |  |
|     |     |     | 340 |     |     |     |     | 345 |     |     |     |     | 350 |     |     |  |  |
| Lys | Pro | Arg | Thr | Asn | Asn | Arg | Thr | Ser | Pro | Arg | Asp | Asn | Thr | Leu | Leu |  |  |
|     |     | 355 |     |     |     |     | 360 |     |     |     |     | 365 |     |     |     |  |  |
| Gln | Gln | Lys | Leu | Leu | Gln | Glu | Ala | Tyr | Met | Thr | Pro | Lys | Asp | Asp | Ile |  |  |
|     | 370 |     |     |     |     | 375 |     |     |     |     | 380 |     |     |     |     |  |  |
| Arg | Leu | Val | Gly | Glu | Leu | Val | Thr | Val | Ile | Gly | Ala | Ile | Ile | Ile | Leu |  |  |
| 385 |     |     |     |     | 390 |     |     |     |     | 395 |     |     |     |     | 400 |  |  |
| Leu | Val | Glu | Val | Pro | Asp | Ile | Phe | Arg | Met | Gly | Val | Thr | Arg | Phe | Phe |  |  |
|     |     |     |     | 405 |     |     |     |     | 410 |     |     |     |     | 415 |     |  |  |
| Gly | Gln | Thr | Ile | Leu | Gly | Gly | Pro | Phe | His | Val | Leu | Ile | Ile | Thr | Tyr |  |  |
|     |     |     | 420 |     |     |     |     | 425 |     |     |     |     | 430 |     |     |  |  |
| Ala | Phe | Met | Val | Leu | Val | Thr | Met | Val | Met | Arg | Leu | Ile | Ser | Ala | Ser |  |  |
|     |     | 435 |     |     |     |     | 440 |     |     |     |     | 445 |     |     |     |  |  |
| Gly | Glu | Val | Val | Pro | Met | Ser | Phe | Ala | Leu | Val | Leu | Gly | Trp | Cys | Asn |  |  |
|     | 450 |     |     |     |     | 455 |     |     |     |     | 460 |     |     |     |     |  |  |
| Val | Met | Tyr | Phe | Ala | Arg | Gly | Phe | Gln | Met | Leu | Gly | Pro | Phe | Thr | Ile |  |  |
| 465 |     |     |     |     | 470 |     |     |     |     | 475 |     |     |     |     | 480 |  |  |
| Met | Asn | Gln | Lys | Met | Ile | Phe | Gly | Asp | Leu | Met | Arg | Phe | Cys | Trp | Leu |  |  |
|     |     |     |     | 485 |     |     |     |     | 490 |     |     |     |     | 495 |     |  |  |
| Met | Ala | Val | Val | Ile | Leu | Gly | Phe | Ala | Ser | Ala | Phe | Tyr | Ile | Ile | Phe |  |  |
|     |     |     | 500 |     |     |     |     | 505 |     |     |     |     | 510 |     |     |  |  |
| Gln | Thr | Glu | Asp | Pro | Glu | Glu | Leu | Gly | His | Phe | Tyr | Asp | Tyr | Pro | Met |  |  |
|     |     | 515 |     |     |     |     | 520 |     |     |     |     | 525 |     |     |     |  |  |
| Ala | Leu | Phe | Ser | Thr | Phe | Glu | Leu | Phe | Leu | Thr | Ile | Ile | Asp | Gly | Pro |  |  |
|     | 530 |     |     |     |     | 535 |     |     |     |     | 540 |     |     |     |     |  |  |
| Ala | Asn | Tyr | Asn | Val | Asp | Leu | Pro | Phe | Met | Tyr | Ser | Ile | Thr | Tyr | Ala |  |  |
| 545 |     |     |     |     | 550 |     |     |     |     | 555 |     |     |     |     | 560 |  |  |
| Ala | Phe | Ala | Ile | Ile | Ala | Thr | Leu | Leu | Met | Leu | Asn | Leu | Leu | Ile | Ala |  |  |
|     |     |     |     | 565 |     |     |     |     | 570 |     |     |     |     | 575 |     |  |  |
| Met | Met | Gly | Asp | Thr | His | Trp | Arg | Val | Ala | His | Glu | Arg | Asp | Glu | Leu |  |  |
|     |     |     | 580 |     |     |     |     | 585 |     |     |     |     | 590 |     |     |  |  |
| Trp | Arg | Ala | Gln | Ile | Val | Ala | Thr | Thr | Val | Met | Leu | Glu | Arg | Lys | Leu |  |  |
|     |     | 595 |     |     |     |     | 600 |     |     |     |     | 605 |     |     |     |  |  |
| Pro | Arg | Cys | Leu | Trp | Pro | Arg | Ser | Gly | Ile | Cys | Gly | Arg | Glu | Tyr | Gly |  |  |

|   |     |     |     |     |
|---|-----|-----|-----|-----|
| 610   |     | 615 |     | 620 |
| Leu Gly Asp Arg Trp Phe Leu Arg Val Glu Asp Arg Gln Asp Leu Asn |     |     |     |     |
| 625   |     | 630 |     | 635 |
| Arg Gln Arg Ile Gln Arg Tyr Ala Gln Ala Phe His Thr Arg Gly Ser |     |     |     |     |
|   | 645 |     | 650 | 655 |
| Glu Asp Leu Asp Lys Asp Ser Val Glu Lys Leu Glu Leu Gly Cys Pro |     |     |     |     |
|   | 660 |     | 665 | 670 |
| Phe Ser Pro His Leu Ser Leu Pro Met Pro Ser Val Ser Arg Ser Thr |     |     |     |     |
|   | 675 |     | 680 | 685 |
| Ser Arg Ser Ser Ala Asn Trp Glu Arg Leu Arg Gln Gly Thr Leu Arg |     |     |     |     |
|   | 690 |     | 695 | 700 |
| Arg Asp Leu Arg Gly Ile Ile Asn Arg Gly Leu Glu Asp Gly Glu Ser |     |     |     |     |
| 705   |     | 710 |     | 715 |
| Trp Glu Tyr Gln Ile   |     |     |     |     |
|   | 725 |     |     |     |

&lt;210&gt; 3

&lt;211&gt; 2175

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 3

```

atgggtttgt cactgcccaa ggagaaaggg ctaattctct gcctatggag caagttctgc 60
agatggttcc agagacggga gtccctgggcc cagagccgag atgagcagaa cctgctgcag 120
cagaagagga tctgggagtc tcctctcctt ctagctgcca aagataatga tgtccaggcc 180
ctgaacaagt tgctcaagta tgaggattgc aaggtgcacc agagaggagc catgggggaa 240
acagcgctac acatagcagc cctctatgac aacctggagg ccgccatggt gctgatggag 300
gctgccccgg agctggtctt tgagcccatg acatctgagc tctatgaggg tcagactgca 360
ctgcacatcg ctgttgtgaa ccagaacatg aacctggtgc gagccctgct tgcccgcagg 420
gccagtgtct ctgccagagc cacaggcact gccttcgcc gtagtccttg caacctcatc 480
tactttgggg agcacccttt gtcctttgct gcctgtgtga acagtgagga gatcgtgcgg 540
ctgctcattg agcatggagc tgacatccgg gccaggact ccctgggaaa cacagtgtta 600
cacatcctca tcctccagcc caacaaaacc ttgacctgcc agatgtacaa cctgttgctg 660
tcctacgaca gacatgggga ccacctgcag cccctggacc tcgtgcccaa tcaccagggt 720
ctcaccctt tcaagctggc tggagtggag ggtaacactg tgatgtttca gcacctgatg 780
cagaagcgga agcacacca gtggacgtat ggaccactga cctcgactct ctatgacctc 840
acagagatcg actcctcagg ggatgagcag tcctgctgg aacttatcat caccaccaag 900
aagcgggagg ctgccagat cctggaccag acgccggtga aggagctggt gaggctcaag 960

```

```
<210> 4
<211> 3245
<212> DNA
<213> Homo sapiens
```

```
<220>
<221> CDS
<222> (84) .. (2699)
```

```
<400> 4
tcgacccacg cgctccgggat ctcccggccg ccggcgccca gccgtcccgg aggctgagca 60
gtgcagacgg gcctggggca ggc atg gcg gat tcc agc gaa ggc ccc cgc gcg 113
                        Met Ala Asp Ser Ser Glu Gly Pro Arg Ala
                        1             5             10
ggg ccc ggg gag gtg gct gag ctg ccc ggg gat gag agt ggc acc cca 161
```

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Gly | Pro | Gly | Glu | Val | Ala | Glu | Leu | Pro | Gly | Asp | Glu | Ser | Gly | Thr | Pro |     |
|     |     |     |     | 15  |     |     |     |     | 20  |     |     |     |     | 25  |     |     |
| ggt | ggg | gag | gct | ttt | cct | ctc | tcc | tcc | ctg | gcc | aat | ctg | ttt | gag | ggg | 209 |
| Gly | Gly | Glu | Ala | Phe | Pro | Leu | Ser | Ser | Leu | Ala | Asn | Leu | Phe | Glu | Gly |     |
|     |     |     | 30  |     |     |     |     | 35  |     |     |     |     | 40  |     |     |     |
| gag | gat | ggc | tcc | ctt | tcg | ccc | tca | ccg | gct | gat | gcc | agt | cgc | cct | gct | 257 |
| Glu | Asp | Gly | Ser | Leu | Ser | Pro | Ser | Pro | Ala | Asp | Ala | Ser | Arg | Pro | Ala |     |
|     |     | 45  |     |     |     |     | 50  |     |     |     |     | 55  |     |     |     |     |
| ggc | cca | ggc | gat | ggg | cga | cca | aat | ctg | cgc | atg | aag | ttc | cag | ggc | gcc | 305 |
| Gly | Pro | Gly | Asp | Gly | Arg | Pro | Asn | Leu | Arg | Met | Lys | Phe | Gln | Gly | Ala |     |
|     | 60  |     |     |     |     | 65  |     |     |     |     | 70  |     |     |     |     |     |
| ttc | cgc | aag | ggg | gtg | ccc | aac | ccc | atc | gat | ctg | ctg | gag | tcc | acc | cta | 353 |
| Phe | Arg | Lys | Gly | Val | Pro | Asn | Pro | Ile | Asp | Leu | Leu | Glu | Ser | Thr | Leu |     |
|     | 75  |     |     |     | 80  |     |     |     |     | 85  |     |     |     |     | 90  |     |
| tat | gag | tcc | tcg | gtg | gtg | cct | ggg | ccc | aag | aaa | gca | ccc | atg | gac | tca | 401 |
| Tyr | Glu | Ser | Ser | Val | Val | Pro | Gly | Pro | Lys | Lys | Ala | Pro | Met | Asp | Ser |     |
|     |     |     |     | 95  |     |     |     |     | 100 |     |     |     |     | 105 |     |     |
| ctg | ttt | gac | tac | ggc | acc | tat | cgt | cac | cac | tcc | agt | gac | aac | aag | agg | 449 |
| Leu | Phe | Asp | Tyr | Gly | Thr | Tyr | Arg | His | His | Ser | Ser | Asp | Asn | Lys | Arg |     |
|     |     |     | 110 |     |     |     |     | 115 |     |     |     |     | 120 |     |     |     |
| tgg | agg | aag | aag | atc | ata | gag | aag | cag | ccg | cag | agc | ccc | aaa | gcc | cct | 497 |
| Trp | Arg | Lys | Lys | Ile | Ile | Glu | Lys | Gln | Pro | Gln | Ser | Pro | Lys | Ala | Pro |     |
|     |     | 125 |     |     |     |     | 130 |     |     |     |     | 135 |     |     |     |     |
| gcc | cct | cag | ccg | ccc | ccc | atc | ctc | aaa | gtc | ttc | aac | cgg | cct | atc | ctc | 545 |
| Ala | Pro | Gln | Pro | Pro | Pro | Ile | Leu | Lys | Val | Phe | Asn | Arg | Pro | Ile | Leu |     |
|     | 140 |     |     |     |     | 145 |     |     |     |     | 150 |     |     |     |     |     |
| ttt | gac | atc | gtg | tcc | cgg | ggc | tcc | act | gct | gac | ctg | gac | ggg | ctg | ctc | 593 |
| Phe | Asp | Ile | Val | Ser | Arg | Gly | Ser | Thr | Ala | Asp | Leu | Asp | Gly | Leu | Leu |     |
|     | 155 |     |     |     | 160 |     |     |     |     | 165 |     |     |     |     | 170 |     |
| cca | ttc | ttg | ctg | acc | cac | aag | aaa | cgc | cta | act | gat | gag | gag | ttt | cga | 641 |
| Pro | Phe | Leu | Leu | Thr | His | Lys | Lys | Arg | Leu | Thr | Asp | Glu | Glu | Phe | Arg |     |
|     |     |     |     | 175 |     |     |     |     | 180 |     |     |     |     | 185 |     |     |
| gag | cca | tct | acg | ggg | aag | acc | tgc | ctg | ccc | aag | gcc | ttg | ctg | aac | ctg | 689 |
| Glu | Pro | Ser | Thr | Gly | Lys | Thr | Cys | Leu | Pro | Lys | Ala | Leu | Leu | Asn | Leu |     |
|     |     |     | 190 |     |     |     |     | 195 |     |     |     |     | 200 |     |     |     |
| agc | aat | ggc | cgc | aac | gac | acc | atc | cct | gtg | ctg | ctg | gac | atc | gcg | gag | 737 |
| Ser | Asn | Gly | Arg | Asn | Asp | Thr | Ile | Pro | Val | Leu | Leu | Asp | Ile | Ala | Glu |     |
|     |     | 205 |     |     |     |     | 210 |     |     |     |     | 215 |     |     |     |     |
| cgc | acc | ggc | aac | atg | agg | gag | ttc | att | aac | tcg | ccc | ttc | cgt | gac | atc | 785 |
| Arg | Thr | Gly | Asn | Met | Arg | Glu | Phe | Ile | Asn | Ser | Pro | Phe | Arg | Asp | Ile |     |
|     | 220 |     |     |     |     | 225 |     |     |     |     | 230 |     |     |     |     |     |
| tac | tat | cga | ggt | cag | aca | gcc | ctg | cac | atc | gcc | att | gag | cgt | cgc | tgc | 833 |
| Tyr | Tyr | Arg | Gly | Gln | Thr | Ala | Leu | His | Ile | Ala | Ile | Glu | Arg | Arg | Cys |     |
|     | 235 |     |     |     | 240 |     |     |     |     | 245 |     |     |     |     | 250 |     |
| aaa | cac | tac | gtg | gaa | ctt | ctc | gtg | gcc | cag | gga | gct | gat | gtc | cac | gcc | 881 |
| Lys | His | Tyr | Val | Glu | Leu | Leu | Val | Ala | Gln | Gly | Ala | Asp | Val | His | Ala |     |

10/17

|   |      |
|---|------|
| gac tac ctg cgg ctg gct ggc gag gtc att acg ctc ttc act ggg gtc | 1649 |
| Asp Tyr Leu Arg Leu Ala Gly Glu Val Ile Thr Leu Phe Thr Gly Val |      |
| 510 515 520   |      |
| ctg ttc ttc ttc acc aac atc aaa gac ttg ttc atg aag aaa tgc cct | 1697 |
| Leu Phe Phe Phe Thr Asn Ile Lys Asp Leu Phe Met Lys Lys Cys Pro |      |
| 525 530 535   |      |
| gga gtg aat tct ctc ttc att gat ggc tcc ttc cag ctg ctc tac ttc | 1745 |
| Gly Val Asn Ser Leu Phe Ile Asp Gly Ser Phe Gln Leu Leu Tyr Phe |      |
| 540 545 550   |      |
| atc tac tct gtc ctg gtg atc gtc tca gca gcc ctc tac ctg gca ggg | 1793 |
| Ile Tyr Ser Val Leu Val Ile Val Ser Ala Ala Leu Tyr Leu Ala Gly |      |
| 555 560 565 570   |      |
| atc gag gcc tac ctg gcc gtg atg gtc ttt gcc ctg gtc ctg ggc tgg | 1841 |
| Ile Glu Ala Tyr Leu Ala Val Met Val Phe Ala Leu Val Leu Gly Trp |      |
| 575 580 585   |      |
| atg aat gcc ctt tac ttc acc cgt ggg ctg aag ctg acg ggg acc tat | 1889 |
| Met Asn Ala Leu Tyr Phe Thr Arg Gly Leu Lys Leu Thr Gly Thr Tyr |      |
| 590 595 600   |      |
| agc atc atg atc cag aag att ctc ttc aag gac ctt ttc cga ttc ctg | 1937 |
| Ser Ile Met Ile Gln Lys Ile Leu Phe Lys Asp Leu Phe Arg Phe Leu |      |
| 605 610 615   |      |
| ctc gtc tac ttg ctc ttc atg atc ggc tac gct tca gcc ctg gtc tcc | 1985 |
| Leu Val Tyr Leu Leu Phe Met Ile Gly Tyr Ala Ser Ala Leu Val Ser |      |
| 620 625 630   |      |
| ctc ctg aac ccg tgt gcc aac atg aag gtg tgc aat gag gac cag acc | 2033 |
| Leu Leu Asn Pro Cys Ala Asn Met Lys Val Cys Asn Glu Asp Gln Thr |      |
| 635 640 645 650   |      |
| aac tgc aca gtg ccc act tac ccc tcg tgc cgt gac agc gag acc ttc | 2081 |
| Asn Cys Thr Val Pro Thr Tyr Pro Ser Cys Arg Asp Ser Glu Thr Phe |      |
| 655 660 665   |      |
| agc acc ttc ctc ctg gac ctg ttt aag ctg acc atc ggc atg ggc gac | 2129 |
| Ser Thr Phe Leu Leu Asp Leu Phe Lys Leu Thr Ile Gly Met Gly Asp |      |
| 670 675 680   |      |
| ctg gag atg ctg agc agc acc aag tac ccc gtg gtc ttc atc atc ctg | 2177 |
| Leu Glu Met Leu Ser Ser Thr Lys Tyr Pro Val Val Phe Ile Ile Leu |      |
| 685 690 695   |      |
| ctg gtg acc tac atc atc ctc acc ttt gtg ctg ctc ctc aac atg ctc | 2225 |
| Leu Val Thr Tyr Ile Ile Leu Thr Phe Val Leu Leu Leu Asn Met Leu |      |
| 700 705 710   |      |
| att gcc ctc atg ggc gag aca gtg ggc cag gtc tcc aag gag agc aag | 2273 |
| Ile Ala Leu Met Gly Glu Thr Val Gly Gln Val Ser Lys Glu Ser Lys |      |
| 715 720 725 730   |      |
| cac atc tgg aag ctg cag tgg gcc acc acc atc ctg gac att gag cgc | 2321 |
| His Ile Trp Lys Leu Gln Trp Ala Thr Thr Ile Leu Asp Ile Glu Arg |      |
| 735 740 745   |      |

tcc ttc ccc gta ttc ctg agg aag gcc ttc cgc tct ggg gag atg gtc 2369  
 Ser Phe Pro Val Phe Leu Arg Lys Ala Phe Arg Ser Gly Glu Met Val  
                     750                    755                    760

acc gtg ggc aag agc tcg gac ggc act cct gac cgc agg tgg tgc ttc 2417  
 Thr Val Gly Lys Ser Ser Asp Gly Thr Pro Asp Arg Arg Trp Cys Phe  
                     765                    770                    775

agg gtg gat gag gtg aac tgg tct cac tgg aac cag aac ttg ggc atc 2465  
 Arg Val Asp Glu Val Asn Trp Ser His Trp Asn Gln Asn Leu Gly Ile  
                     780                    785                    790

atc aac gag gac ccg ggc aag aat gag acc tac cag tat tat ggc ttc 2513  
 Ile Asn Glu Asp Pro Gly Lys Asn Glu Thr Tyr Gln Tyr Tyr Gly Phe  
                     795                    800                    805                    810

tcg cat acc gtg ggc cgc ctc cgc agg gat cgc tgg tcc tcg gtg gta 2561  
 Ser His Thr Val Gly Arg Leu Arg Arg Asp Arg Trp Ser Ser Val Val  
                     815                    820                    825

ccc cgc gtg gtg gaa ctg aac aag aac tcg aac ccg gac gag gtg gtg 2609  
 Pro Arg Val Val Glu Leu Asn Lys Asn Ser Asn Pro Asp Glu Val Val  
                     830                    835                    840

gtg cct ctg gac agc atg ggg aac ccc cgc tgc gat ggc cac cag cag 2657  
 Val Pro Leu Asp Ser Met Gly Asn Pro Arg Cys Asp Gly His Gln Gln  
                     845                    850                    855

ggt tac ccc cgc aag tgg agg act gat gac gcc ccg ctc tag 2699  
 Gly Tyr Pro Arg Lys Trp Arg Thr Asp Asp Ala Pro Leu  
                     860                    865                    870

ggactgcagc ccagccccag ottctctgcc cactcatttc tagtccagcc gcatttcagc 2759

agtgccttct ggggtgtccc cccacaccct gctttggccc cagaggcgag ggaccagtgg 2819

aggtgccagg gaggccccag gaccctgtgg tcccctggct ctgcctcccc accctgggggt 2879

ggggggtccc ggccacctgt cttgctccta tggagtcaca taagccaacg ccagagcccc 2939

tccacctcag gccccagccc ctgcctctcc attatttatt tgctctgctc tcaggaagcg 2999

acgtgacccc tgccccagct ggaacctggc agaggcctta ggaccccggt ccaagtgcac 3059

tgcccggcca agccccagcc tcagcctgcg cctgagctgc atgcgccacc atttttggca 3119

gcgtggcagc tttgcaagg gctggggccc tcggcggtgg gccatgcctt ctgtgtgttc 3179

tgtagtgtct gggatttgcc ggtgctcaat aaatggttat tcattgaaaa aaaaaaaaaa 3239

aaaagg 3245

<210> 5

<211> 871

<212> PRT

<213> Homo sapiens

<400> 5

Met Ala Asp Ser Ser Glu Gly Pro Arg Ala Gly Pro Gly Glu Val Ala  
           1                    5                    10                    15

Glu Leu Pro Gly Asp Glu Ser Gly Thr Pro Gly Gly Glu Ala Phe Pro  
                   20                                  25                                  30  
 Leu Ser Ser Leu Ala Asn Leu Phe Glu Gly Glu Asp Gly Ser Leu Ser  
                   35                                  40                                  45  
 Pro Ser Pro Ala Asp Ala Ser Arg Pro Ala Gly Pro Gly Asp Gly Arg  
                   50                                  55                                  60  
 Pro Asn Leu Arg Met Lys Phe Gln Gly Ala Phe Arg Lys Gly Val Pro  
                   65                                  70                                  75                                  80  
 Asn Pro Ile Asp Leu Leu Glu Ser Thr Leu Tyr Glu Ser Ser Val Val  
                                   85                                  90                                  95  
 Pro Gly Pro Lys Lys Ala Pro Met Asp Ser Leu Phe Asp Tyr Gly Thr  
                                   100                                  105                                  110  
 Tyr Arg His His Ser Ser Asp Asn Lys Arg Trp Arg Lys Lys Ile Ile  
                   115                                  120                                  125  
 Glu Lys Gln Pro Gln Ser Pro Lys Ala Pro Ala Pro Gln Pro Pro Pro  
                   130                                  135                                  140  
 Ile Leu Lys Val Phe Asn Arg Pro Ile Leu Phe Asp Ile Val Ser Arg  
                   145                                  150                                  155                                  160  
 Gly Ser Thr Ala Asp Leu Asp Gly Leu Leu Pro Phe Leu Leu Thr His  
                                   165                                  170                                  175  
 Lys Lys Arg Leu Thr Asp Glu Glu Phe Arg Glu Pro Ser Thr Gly Lys  
                                   180                                  185                                  190  
 Thr Cys Leu Pro Lys Ala Leu Leu Asn Leu Ser Asn Gly Arg Asn Asp  
                   195                                  200                                  205  
 Thr Ile Pro Val Leu Leu Asp Ile Ala Glu Arg Thr Gly Asn Met Arg  
                   210                                  215                                  220  
 Glu Phe Ile Asn Ser Pro Phe Arg Asp Ile Tyr Tyr Arg Gly Gln Thr  
                   225                                  230                                  235                                  240  
 Ala Leu His Ile Ala Ile Glu Arg Arg Cys Lys His Tyr Val Glu Leu  
                                   245                                  250                                  255  
 Leu Val Ala Gln Gly Ala Asp Val His Ala Gln Ala Arg Gly Arg Phe  
                                   260                                  265                                  270  
 Phe Gln Pro Lys Asp Glu Gly Gly Tyr Phe Tyr Phe Gly Glu Leu Pro  
                   275                                  280                                  285  
 Leu Ser Leu Ala Ala Cys Thr Asn Gln Pro His Ile Val Asn Tyr Leu  
                   290                                  295                                  300  
 Thr Glu Asn Pro His Lys Lys Ala Asp Met Arg Arg Gln Asp Ser Arg  
                   305                                  310                                  315                                  320  
 Gly Asn Thr Val Leu His Ala Leu Val Ala Ile Ala Asp Asn Thr Arg  
                                   325                                  330                                  335



Glu Asn Thr Lys Phe Val Thr Lys Met Tyr Asp Leu Leu Leu Leu Lys  
 340 345 350  
 Cys Ala Arg Leu Phe Pro Asp Ser Asn Leu Glu Ala Val Leu Asn Asn  
 355 360 365  
 Asp Gly Leu Ser Pro Leu Met Met Ala Ala Lys Thr Gly Lys Ile Gly  
 370 375 380  
 Ile Phe Gln His Ile Ile Arg Arg Glu Val Thr Asp Glu Asp Thr Arg  
 385 390 395 400  
 His Leu Ser Arg Lys Phe Lys Asp Trp Ala Tyr Gly Pro Val Tyr Ser  
 405 410 415  
 Ser Leu Tyr Asp Leu Ser Ser Leu Asp Thr Cys Gly Glu Glu Ala Ser  
 420 425 430  
 Val Leu Glu Ile Leu Val Tyr Asn Ser Lys Ile Glu Asn Arg His Glu  
 435 440 445  
 Met Leu Ala Val Glu Pro Ile Asn Glu Leu Leu Arg Asp Lys Trp Arg  
 450 455 460  
 Lys Phe Gly Ala Val Ser Phe Tyr Ile Asn Val Val Ser Tyr Leu Cys  
 465 470 475 480  
 Ala Met Val Ile Phe Thr Leu Thr Ala Tyr Tyr Gln Pro Leu Glu Gly  
 485 490 495  
 Thr Pro Pro Tyr Pro Tyr Arg Thr Thr Val Asp Tyr Leu Arg Leu Ala  
 500 505 510  
 Gly Glu Val Ile Thr Leu Phe Thr Gly Val Leu Phe Phe Phe Thr Asn  
 515 520 525  
 Ile Lys Asp Leu Phe Met Lys Lys Cys Pro Gly Val Asn Ser Leu Phe  
 530 535 540  
 Ile Asp Gly Ser Phe Gln Leu Leu Tyr Phe Ile Tyr Ser Val Leu Val  
 545 550 555 560  
 Ile Val Ser Ala Ala Leu Tyr Leu Ala Gly Ile Glu Ala Tyr Leu Ala  
 565 570 575  
 Val Met Val Phe Ala Leu Val Leu Gly Trp Met Asn Ala Leu Tyr Phe  
 580 585 590  
 Thr Arg Gly Leu Lys Leu Thr Gly Thr Tyr Ser Ile Met Ile Gln Lys  
 595 600 605  
 Ile Leu Phe Lys Asp Leu Phe Arg Phe Leu Leu Val Tyr Leu Leu Phe  
 610 615 620  
 Met Ile Gly Tyr Ala Ser Ala Leu Val Ser Leu Leu Asn Pro Cys Ala  
 625 630 635 640  
 Asn Met Lys Val Cys Asn Glu Asp Gln Thr Asn Cys Thr Val Pro Thr  
 645 650 655  
 Tyr Pro Ser Cys Arg Asp Ser Glu Thr Phe Ser Thr Phe Leu Leu Asp

| 660 |     |     |     |     | 665 |     |     |     |     | 670 |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Leu | Phe | Lys | Leu | Thr | Ile | Gly | Met | Gly | Asp | Leu | Glu | Met | Leu | Ser | Ser |
|     |     | 675 |     |     |     |     | 680 |     |     |     |     | 685 |     |     |     |
| Thr | Lys | Tyr | Pro | Val | Val | Phe | Ile | Ile | Leu | Leu | Val | Thr | Tyr | Ile | Ile |
|     | 690 |     |     |     |     | 695 |     |     |     |     | 700 |     |     |     |     |
| Leu | Thr | Phe | Val | Leu | Leu | Leu | Asn | Met | Leu | Ile | Ala | Leu | Met | Gly | Glu |
| 705 |     |     |     |     | 710 |     |     |     |     | 715 |     |     |     |     | 720 |
| Thr | Val | Gly | Gln | Val | Ser | Lys | Glu | Ser | Lys | His | Ile | Trp | Lys | Leu | Gln |
|     |     |     |     | 725 |     |     |     |     | 730 |     |     |     |     | 735 |     |
| Trp | Ala | Thr | Thr | Ile | Leu | Asp | Ile | Glu | Arg | Ser | Phe | Pro | Val | Phe | Leu |
|     |     |     |     | 740 |     |     |     | 745 |     |     |     |     | 750 |     |     |
| Arg | Lys | Ala | Phe | Arg | Ser | Gly | Glu | Met | Val | Thr | Val | Gly | Lys | Ser | Ser |
|     |     | 755 |     |     |     |     | 760 |     |     |     |     | 765 |     |     |     |
| Asp | Gly | Thr | Pro | Asp | Arg | Arg | Trp | Cys | Phe | Arg | Val | Asp | Glu | Val | Asn |
|     | 770 |     |     |     |     | 775 |     |     |     |     | 780 |     |     |     |     |
| Trp | Ser | His | Trp | Asn | Gln | Asn | Leu | Gly | Ile | Ile | Asn | Glu | Asp | Pro | Gly |
| 785 |     |     |     |     | 790 |     |     |     |     | 795 |     |     |     |     | 800 |
| Lys | Asn | Glu | Thr | Tyr | Gln | Tyr | Tyr | Gly | Phe | Ser | His | Thr | Val | Gly | Arg |
|     |     |     |     | 805 |     |     |     |     | 810 |     |     |     |     | 815 |     |
| Leu | Arg | Arg | Asp | Arg | Trp | Ser | Ser | Val | Val | Pro | Arg | Val | Val | Glu | Leu |
|     |     |     | 820 |     |     |     |     | 825 |     |     |     |     | 830 |     |     |
| Asn | Lys | Asn | Ser | Asn | Pro | Asp | Glu | Val | Val | Val | Pro | Leu | Asp | Ser | Met |
|     |     | 835 |     |     |     |     | 840 |     |     |     |     | 845 |     |     |     |
| Gly | Asn | Pro | Arg | Cys | Asp | Gly | His | Gln | Gln | Gly | Tyr | Pro | Arg | Lys | Trp |
|     | 850 |     |     |     |     | 855 |     |     |     |     | 860 |     |     |     |     |
| Arg | Thr | Asp | Asp | Ala | Pro | Leu |     |     |     |     |     |     |     |     |     |
| 865 |     |     |     |     | 870 |     |     |     |     |     |     |     |     |     |     |

&lt;210&gt; 6

&lt;211&gt; 2613

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 6

```

atggcggatt ccagcgaagg cccccgcgcg gggcccgggg aggtggctga gctccccggg 60
gatgagagtg gcaccccagg tggggaggct tttcctctct cctccctggc caatctgttt 120
gagggggagg atggctccct ttcgccctca ccggctgatg ccagtcgccc tgctggccca 180
ggcgatgggc gaccaaactc gcgcatgaag ttccagggcg ccttcogcaa gggggtgccc 240
aaccatcatc atctgctgga gtccacccta tatgagtcct cgggtggtgcc tgggccaag 300
aaagcaccca tggactcact gtttgactac ggcacctatc gtcaccactc cagtgacaac 360
aagaggtgga ggaagaagat catagagaag cagccgcaga gcccacaaag ccctgccct 420

```

cagccgcccc ccacccctcaa agtcttcaac cggcctatcc tctttgacat cgtgtcccgg 480  
ggctccactg ctgacctgga cgggctgctc ccattcttgc tgaccacaaa gaaacgccta 540  
actgatgagg agtttcgaga gccatctacg gggaagacct gcctgcccga ggcccttgcgtg 600  
aacctgagca atggccgcaa cgacaccatc cctgtgctgc tggacatcgc ggagcgcacc 660  
ggcaacatga gggagttcat taactcgccc ttccgtgaca tctactatcg aggtcagaca 720  
gccctgcaca tcgccattga gcgtcgctgc aaacactacg tggaaacttct cgtggcccag 780  
ggagctgatg tccacgcccc ggcccgtggg cgcttcttcc agcccaggga tgaggggggc 840  
tacttctact ttggggagct gcccctgtcg ctggctgcct gcaccaacca gcccacatt 900  
gtcaactacc tgacggagaa cccccacaag aaggcggaca tgcggcgcca ggactcgcga 960  
ggcaacacag tgctgcatgc gctggtggcc attgctgaca acaccctga gaacaccaag 1020  
tttgttacca agatgtacga cctgctgctg ctcaagtgtg cccgcctctt ccccgacagc 1080  
aacctggagg ccgtgctcaa caacgacggc ctctcgcccc tcatgatggc tgccaagacg 1140  
ggcaagattg ggatctttca gcacatcatc cggcgggagg tgacggatga ggacacacgg 1200  
cacctgtccc gcaagttcaa ggactgggcc tatgggccag tgtattcctc gctttatgac 1260  
ctctcctccc tggacacgtg tggggaagag gcctccgtgc tggagatcct ggtgtacaac 1320  
agcaagattg agaaccgcca cgagatgctg gctgtggagc ccatcaatga actgctgcgg 1380  
gacaagtggc gcaagttcgg ggccgtctcc ttctacatca acgtggtctc ctacctgtgt 1440  
gcatgttca tcttactctt caccgcctac taccagccgc tggagggcac accgccgtac 1500  
ccttaccgca ccacggtgga ctacctgcgg ctggctggcg aggtcattac gctcttact 1560  
ggggtcctgt tcttcttcac caacatcaaa gacttgttca tgaagaaatg ccctggagtgt 1620  
aattctctct tcattgatgg ctccctccag ctgctctact tcatctactc tgtcctggtg 1680  
atcgtctcag cagccctcta cctggcaggg atcgaggcct acctggccgt gatggtcttt 1740  
gccctggtcc tgggctggat gaatgccctt tacttcaccc gtgggctgaa gctgacgggg 1800  
acctatagca tcatgatcca gaagattctc ttcaaggacc ttttccgatt cctgctcgtc 1860  
tacttgctct tcatgatcgg ctacgcttca gccctggtct ccctcctgaa cccgtgtgcc 1920  
aacatgaagg tgtgcaatga ggaccagacc aactgcacag tgcccactta cccctcgtgc 1980  
cgtgacagcg agaccttcag caccctcctc ctggacctgt ttaagctgac catcggcagt 2040  
ggcgacctgg agatgctgag cagcaccaag taccctgtgg tcttcatcat cctgctggtg 2100  
acctacatca tctcacctt tgtgctgctc ctcaacatgc tcattgccct catgggagag 2160  
acagtgggcc aggtctccaa ggagagcaag cacatctgga agctgcagtg ggccaccacc 2220

atcctggaca ttgagcgctc cttccccgta ttcctgagga aggccttccg ctctggggag 2280  
atggtcaccg tgggcaagag ctgggacggc actcctgacc gcaggtggtg cttcaggggtg 2340  
gatgaggtga actggtctca ctggaaccag aacttgggca tcatcaacga ggacccgggc 2400  
aagaatgaga cctaccagta ttatggcttc tcgcataccg tgggcgcgct ccgcagggat 2460  
cgctggtcct cggtggtacc ccgcgtggtg gaactgaaca agaactcgaa cccggacgag 2520  
gtggtggtgc ctctggacag catggggaac ccccgctgcg atggccacca gcagggttac 2580  
ccccgcaagt ggaggactga tgacgccccg ctc 2613

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 01/08329

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/12 C12N15/62 C12N15/63 C07K14/705 C07K16/28  
 C12Q1/68 G01N33/577 G01N33/68

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N C07K C12Q G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, SEQUENCE SEARCH

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category * | Citation of document, with indication, where appropriate, of the relevant passages   | Relevant to claim No.  |
|------------|--|------------------------|
| X          | DE 198 11 194 A (METAGEN GES FUER<br>GENOMFORSCHUN)<br>16 September 1999 (1999-09-16)<br>page 182; claims 1-35<br>SEQ ID NO.48<br>page 111 -page 112<br>---                  | 5-12,<br>20-22         |
| P,X        | WO 01 14423 A (SMITHKLINE BEECHAM PLC)<br>1 March 2001 (2001-03-01)<br><br>SEQ ID NOs.1 and 2<br>page 1, line 21 - line 29<br>page 7, line 20 -page 15, line 4<br>---<br>-/- | 1,2,<br>4-26,<br>29-35 |

☒ Further documents are listed in the continuation of box C.☒ Patent family members are listed in annex.

## \* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance  
 "E" earlier document but published on or after the international filing date  
 "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)  
 "O" document referring to an oral disclosure, use, exhibition or other means  
 "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

23 November 2001

Date of mailing of the international search report

26. 02. 2002

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
 NL - 2280 HV Rijswijk  
 Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
 Fax: (+31-70) 340-3016

Authorized officer

Blanco Urgoiti, B

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 01/08329

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

| Category * | Citation of document, with indication, where appropriate, of the relevant passages  | Relevant to claim No.        |
|------------|---|------------------------------|
| P,X        | WO 01 04303 A (HEDIGER MATTHIAS A)<br>18 January 2001 (2001-01-18)<br><br>SEQ ID NOs.1 and 2<br>page 9, line 4 -page 12, line 26<br>page 16, line 5 -page 17, line 3<br>--- | 1,2,<br>4-22,25,<br>37-40,42 |
| A          | WO 99 37675 A (BRAKE ANTHONY J ;JULIUS<br>DAVID J (US); UNIV CALIFORNIA (US);<br>CATERI) 29 July 1999 (1999-07-29)<br>---   |                              |
| A          | EP 0 943 683 A (SMITHKLINE BEECHAM PLC)<br>22 September 1999 (1999-09-22)<br>-----  |                              |

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US 01/08329**Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)**

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:  
**Although claims 31 to 35 and 37 to 42 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.**
2. ☐ Claims Nos.:  
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

**Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)**

This International Searching Authority found multiple inventions in this international application, as follows:

**see additional sheet**

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

**see FURTHER INFORMATION sheet, invention 1.**

**Remark on Protest**

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

## FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: (1,2,4-43)-partially

The nucleic acid sequences of SEQ ID NOs.1 and/or 3; the polypeptide of SEQ ID NO.2; fragments, allelic variants or nucleic acid molecules hybridizing said molecules; vectors; host cells; antibodies; methods to detect the presence of any of the mentioned polypeptides; methods to detect the presence of any of the mentioned nucleic acids; methods to detect ligands; methods to modulate the activity of the mentioned polypeptides; methods to identify modulators; methods of treatment and methods of identifying a proliferation disorder thereof.

2. Claims: (1,2,4- 43)-partially; 3-complete

The nucleic acid sequences of SEQ ID NOs.4 and/or 6; the polypeptide of SEQ ID NO.5; fragments, allelic variants or nucleic acid molecules hybridizing said molecules; plasmid deposited with ATCC as Acc No. PTA-2013; vectors; host cells; antibodies; methods to detect the presence of any of the mentioned polypeptides; methods to detect the presence of any of the mentioned nucleic acids; methods to detect ligands; methods to modulate the activity of the mentioned polypeptides; methods to identify modulators; methods of treatment and methods of identifying a proliferation disorder thereof.



## INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 01/08329

| Patent document<br>cited in search report |   | Publication<br>date | Patent family<br>member(s) | Publication<br>date |
|---|---|---------------------|----------------------------|---------------------|
| DE 19811194                               | A | 16-09-1999          | DE 19811194 A1             | 16-09-1999          |
|   |   |                     | WO 9946375 A2              | 16-09-1999          |
|   |   |                     | EP 1068318 A2              | 17-01-2001          |
| -----                                     |   |                     |                            |                     |
| WO 0114423                                | A | 01-03-2001          | WO 0114423 A1              | 01-03-2001          |
| -----                                     |   |                     |                            |                     |
| WO 0104303                                | A | 18-01-2001          | AU 5778600 A               | 30-01-2001          |
|   |   |                     | WO 0104303 A1              | 18-01-2001          |
| -----                                     |   |                     |                            |                     |
| WO 9937675                                | A | 29-07-1999          | AU 2466799 A               | 09-08-1999          |
|   |   |                     | AU 742391 B2               | 03-01-2002          |
|   |   |                     | AU 9115698 A               | 08-03-1999          |
|   |   |                     | EP 1009804 A1              | 21-06-2000          |
|   |   |                     | EP 1047711 A1              | 02-11-2000          |
|   |   |                     | JP 2001514879 T            | 18-09-2001          |
|   |   |                     | WO 9909140 A1              | 25-02-1999          |
|   |   |                     | WO 9937675 A1              | 29-07-1999          |
|   |   |                     | US 6335180 B1              | 01-01-2002          |
| -----                                     |   |                     |                            |                     |
| EP 0943683                                | A | 22-09-1999          | EP 0943683 A1              | 22-09-1999          |
|   |   |                     | JP 11279196 A              | 12-10-1999          |
|   |   |                     | US 6239267 B1              | 29-05-2001          |
|   |   |                     | US 2001047090 A1           | 29-11-2001          |
| -----                                     |   |                     |                            |                     |